



# Etude du couplage du boson de Higgs au quark top dans l'expérience ATLAS

Arthur Chomont

## ► To cite this version:

Arthur Chomont. Etude du couplage du boson de Higgs au quark top dans l'expérience ATLAS. High Energy Physics - Experiment [hep-ex]. 2014. in2p3-01011123

**HAL Id: in2p3-01011123**

**<https://hal.in2p3.fr/in2p3-01011123>**

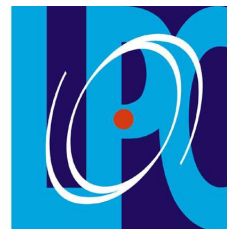
Submitted on 23 Jun 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



UFR Sciences et Technologies



Laboratoire de Physique Corpusculaire  
de Clermont-Ferrand

# **MASTER SCIENCES DE LA MATIERE DEUXIÈME ANNÉE**

**SPÉCIALITÉ : Physique des Particules**

**RAPPORT DE STAGE**

## ***Etude du couplage du boson de Higgs au quark top dans l'expérience ATLAS***

par

**Arthur CHOMONT**

Responsable(s) de stage : **Djamel BOUMEDIENE**



**Juin 2014**



# Remerciements

Je souhaite tout d'abord remercier M.Falvard, directeur du laboratoire de physique corpusculaire pour m'avoir reçu au sein du laboratoire, ainsi que Dominique Pallin, responsable du groupe ATLAS pour m'avoir accueilli au sein de cette équipe.

Je remercie sincèrement Djamel Boumediene, chercheur au CNRS, pour avoir accepté d'encadrer mon stage ainsi que pour sa patience et sa pédagogie.

Je souhaite également remercier Loic Valéry, étudiant en troisième année de thèse et Emmanuelle Dubreuil, étudiante en troisième année de thèse pour avoir répondu avec patience et clarté à mes questions.

Merci aussi à tous les autres membres de l'équipe ATLAS pour m'avoir si bien accueilli durant ce stage.

# Table des matières

<b>Introduction</b>	<b>1</b>
<b>1 Contexte théorique : le Modèle Standard de la physique des particules</b>	<b>2</b>
1.1 Le Modèle Standard et ses limites . . . . .	2
1.1.1 Les particules élémentaires . . . . .	2
1.1.2 Interactions fondamentales . . . . .	3
1.1.3 Les limites du Modèle Standard . . . . .	4
1.2 Phénoménologie du boson de Higgs . . . . .	5
1.2.1 Constante de couplage du boson de Higgs à un fermion . . . . .	5
1.2.2 Les voies de productions et désintégrations du boson de Higgs . . . . .	5
1.2.3 Le choix de l'étude du processus $t\bar{t}H$ . . . . .	6
<b>2 Outils expérimentaux : Le LHC et le détecteur ATLAS</b>	<b>7</b>
2.1 Le LHC et les collisions proton-proton . . . . .	7
2.1.1 L'accélérateur . . . . .	7
2.1.2 Les collisions protons-protons . . . . .	8
2.2 ATLAS un détecteur de physique des particules . . . . .	8
2.2.1 Le détecteur interne . . . . .	8
2.2.2 Les calorimètres . . . . .	9
2.2.3 Le spectromètre à muons . . . . .	10
2.2.4 Le système de déclenchement . . . . .	10
2.2.5 Conventions géométriques . . . . .	10
2.2.6 ATLAS, une collaboration internationale et une équipe à Clermont-Ferrand . . . . .	11
2.3 Les données utilisées dans cette analyse . . . . .	11
<b>3 Signature choisie et préselection pour l'étude du processus <math>t\bar{t}H</math></b>	<b>12</b>
3.1 Les objets reconstruits dans ATLAS . . . . .	12
3.2 Choix de la signature expérimentale : deux leptons de même signe . . . . .	13
3.3 Génération des événements Monte-Carlo . . . . .	14
3.4 Bruit de fond . . . . .	15

3.4.1	Les incertitudes sur les bruits de fond simulés par Monte-Carlo . . . . .	15
3.4.2	Bruits de fond irréductibles . . . . .	15
3.4.3	Les événements <i>Fakes</i> . . . . .	16
3.4.4	Les événements <i>Mis-Id</i> . . . . .	16
3.5	Modélisation du signal et reconstruction des objets hadroniques . . . . .	17
3.6	Définition de la signification statistique . . . . .	17
3.7	Préselection . . . . .	18
<b>4</b>	<b>Estimation et contrôle des bruits de fond</b>	<b>19</b>
4.1	Bruit de fond global . . . . .	19
4.2	Processus $t\bar{t}V$ . . . . .	21
4.3	Estimation du nombre d'événements <i>Fakes</i> depuis les données . . . . .	22
4.3.1	Définition de la méthode . . . . .	22
4.3.2	Incertitude statistique . . . . .	23
4.3.3	Incertitudes systématiques . . . . .	23
4.3.4	Nombres d'événements <i>Fakes</i> obtenus . . . . .	23
<b>5</b>	<b>Résultat et calcul de limite</b>	<b>25</b>
5.1	Définition de la région de signal . . . . .	25
5.2	Définition de la méthode $CL_s$ . . . . .	27
5.2.1	Définition de la grandeur $CL_s$ . . . . .	27
5.2.2	Limite observée . . . . .	28
5.2.3	Définition de la force du signal $\mu$ . . . . .	28
5.2.4	Limite attendue médiane . . . . .	28
5.3	Application à l'analyse du signal $t\bar{t}H$ . . . . .	29
	<b>Conclusion</b>	<b>30</b>
<b>A</b>	<b>Distributions des bruits de fond et des données dans la régions de contrôle</b>	<b>XXXI</b>
	<b>Bibliographie</b>	<b>XXXVI</b>
	<b>Résumé</b>	<b>XXXVII</b>
	<b>Abstract</b>	<b>XXXVII</b>

# Introduction

Le Modèle Standard est à l'heure actuelle le cadre théorique de la physique de l'infiniment petit. De nombreuses expériences ont jusqu'à aujourd'hui confirmé la validité du Modèle Standard malgré la certitude que le Modèle Standard est incomplet et qu'une autre théorie doit être mise en place, qui engloberait le Modèle Standard et comblerait ses lacunes. Les théories s'adressant à ce problème sont regroupés sous l'appellation de Nouvelle Physique.

L'exemple le plus récent de la validité de ce modèle est la découverte en Juillet 2012 au LHC par les expériences ATLAS [1] et CMS [2] du boson de Higgs dont l'existence avait été prédite en 1964 par Robert Brout, François Englert et Peter Higgs. Dans le Modèle Standard cette particule est à l'origine de la masse des particules. Il faut maintenant étudier les caractéristiques de cette particule pour savoir si le boson de Higgs découvert correspond réellement à celui prédit dans le Modèle Standard ou bien à un celui d'un scénario de Nouvelle Physique.

Parmi les propriétés du boson de Higgs à tester, son couplage aux fermions est très intéressant car lié à la masse de ces particules. Le couplage au quark top devrait donc être le plus intense. L'étude du processus  $t\bar{t}H$  est un bon moyen de tester le Modèle Standard par une mesure directe de ce couplage .

Dans un premier temps, après une présentation rapide du Modèle Standard et des appareils qui nous permettent de mener notre étude, le LHC et le détecteur ATLAS, nous présenterons la raison du choix de l'étude du couplage du quark top au boson de Higgs par le biais d'événements avec deux leptons de même signe dans l'état final. Ensuite, une étude plus spécifique sera faite sur les bruits de fond présents dans cette analyse. Enfin nous nous intéresserons à l'estimation de la sensibilité de l'analyse au signal étudié.

# 1 Contexte théorique : le Modèle Standard de la physique des particules

## 1.1 Le Modèle Standard et ses limites

Le Modèle Standard décrit l'ensemble des particules élémentaires. Elles peuvent par exemple être séparées en deux groupes : les fermions et les bosons. A chacune des particules décrites ci-dessous il faut ajouter une antiparticule de même masse mais de charge électrique opposée.

### 1.1.1 Les particules élémentaires

Les **fermions** sont des particules de spin demi-entier qui obéissent à la statistique de Fermi-Dirac et constituent la matière. Ils peuvent être séparés en deux sous-groupes : les **leptons** et les **quarks**. Ces derniers ont la propriété d'être sensibles à l'interaction forte car porteur de charges de couleur (voir la partie 1.2.3), ce qui n'est pas le cas des leptons. Il existe trois familles dans chacun de ces deux sous-groupes, la matière ordinaire étant constituée exclusivement de fermions de la première famille. Les particules de cette famille sont pour les quarks, le quark *up* et le quark *down* qui sont les bases des protons (*uud*) et neutrons (*udd*) et pour les leptons, l'électron et le neutrino électronique. Les autres quarks et leptons de deuxième et troisième génération sont listés dans le Tableau 1.1.

TAB. 1.1: Particules élémentaires

leptons		quarks	
$e^-$ électron	$\nu_e$ neutrino e	u up	d down
$\mu^-$ muon	$\nu_\mu$ neutrino $\mu$	s étrange	c charmé
$\tau^-$ tau	$\nu_\tau$ neutrino $\mu$	b beau	t top

Les **bosons** ont pour leur part un spin entier et obéissent à la statistique de Bose-Einstein. Ces particules portent les interactions entre les différents fermions. Il existe au moins un boson par interaction fondamentale décrite dans la suite du chapitre (excepté pour la gravitation dont l'existence d'un boson lié à cette interaction n'est qu'hypothétique). Ces bosons vecteurs (de spin égal à un) sont appelés les bosons  $W^\pm$ ,  $Z$ , le photon et les gluons. Il existe un autre boson, scalaire (de spin nul), le boson de Higgs, qui représente l'interaction entre les particules et le champ de Higgs, cette interaction donnant une masse aux particules.



### 1.1.2 Interactions fondamentales

Chaque corps ou particule matérielle est soumis à tout instant à de nombreuses forces telles que les forces électromagnétique, gravitationnelle, de frottement (pour un corps en mouvement)... Notre connaissance actuelle de la physique nous permet d'affirmer qu'il existe quatre interactions fondamentales qui peuvent décrire toutes ces forces. Par définition une interaction est l'action réciproque de deux corps l'un sur l'autre. Les quatre interactions fondamentales sont les interactions électromagnétique, gravitationnelle, faible et forte. Chacune d'elles agit sur des fermions qui portent respectivement, une charge électrique, une masse, une charge d'isospin faible ou une charge de couleur.

La portée de chaque interaction fondamentale est liée à la masse des bosons correspondants d'après le modèle de Yukawa. Dans ce cadre, chaque particule soumise à une de ces interactions est dans un potentiel de la forme :

$$U(r) = \frac{C}{r} e^{-mr} \quad (1.1)$$

avec  $C$  une constante,  $r$  la portée de la force et  $m$  la masse du boson qui transporte l'interaction. Ainsi pour un boson de masse nulle, la portée de l'interaction peut être infinie alors que plus un boson va être massif et plus ce potentiel va décroître rapidement et la portée de l'interaction sera donc limitée.

#### L'interaction forte

Les particules sujettes à cette interaction sont les quarks, les bosons vecteurs de cette force sont les **gluons** (au nombre de huit). Ces différentes particules portent toutes des **charges de couleur**, avec une couleur unique pour les quarks et une charge de couleur et d'anti-couleur pour les gluons. Les couleurs sont au nombre de trois : Rouge, Bleu et Vert. A celles-ci ils faut associer trois anti-couleurs et l'association de ces trois couleurs (ou anticouleurs) est neutre (blanche).

Cette interaction est décrite par la chromodynamique quantique (QCD en Anglais) et, malgré la masse nulle des gluons, cette interaction a une portée extrêmement faible (de l'ordre de  $10^{-18}$  m). En effet plus deux quarks en interaction forte l'un avec l'autre vont être éloignés et plus l'énergie nécessaire pour continuer à les séparer va augmenter jusqu'à devenir infinie (pour une distance entre les quarks de l'ordre de  $10^{-18}$  m) : c'est ce que l'on appelle le *confinement des quarks*.

Ainsi, les quarks ne peuvent jamais être trouvés isolés dans la nature et on ne peut donc observer que des baryons (particule formée de trois quarks) et des mésons (association d'un quark et d'un anti-quark). Toutes ces particules n'ont pas de couleur, soit par association des trois couleurs dans les baryons (Rouge+Bleu+Vert=Blanc), soit par combinaison d'une couleur et de son anticouleur dans les mésons (Vert+anti-Vert=Blanc).

Le quark *top* est une exception à cette règle. Son temps de vie est plus court que l'échelle de temps de la *QCD* : il se désintègre en un boson  $W$  et un quark  $b$  avant de s'hadroniser. <sup>1</sup>

#### L'interaction faible

L'interaction faible relie les particules possédant une **charge d'isospin faible**, c'est-à-dire toutes les particules. L'isospin faible est un nombre quantique qui vaut  $1/2$  pour les neutrinos, les quarks  $u, c$  et  $t$  et qui vaut  $-1/2$  pour les leptons chargés et les quarks  $d, s$  et  $b$ . Ce nombre est noté  $I_3$ . L'interaction faible entre en jeu dans le mécanisme de la désintégration  $\beta$  :

$$n \longrightarrow p + e^- + \bar{\nu}_e \quad (1.2)$$

---

<sup>1</sup>Le confinement des quarks imposent qu'ils soient couplés à d'autres quarks. Un quark seul va donc se coupler avec un quark du vide et ce phénomène est appelé hadronisation.

Elle agit sur une très courte distance (de l'ordre de  $10^{-16}$  m) car les bosons vecteurs de cette interaction,  $W^\pm$  et  $Z$  ont des masses importantes ( $80.39 \text{ GeV}/c^2$  [3] pour les deux premiers et  $91.19 \text{ GeV}/c^2$  [3] pour le dernier).

## L'interaction électromagnétique

Cette interaction agit sur les particules portant une **charge électrique** tels que les quarks ou les leptons chargés. Le boson responsable de la propagation de cette interaction est le **photon**. Sa masse étant nulle, l'interaction électromagnétique a une portée infinie. Son intensité est proportionnelle à la charge électrique des particules considérées et inversement proportionnelle à la distance au carré entre les deux particules.

## La symétrie électrofaible et sa brisure

Pour expliquer la désintégration des neutrons par radioactivité  $\beta$ , Salam et Glashow ont imaginé l'unification des deux interactions faible et électromagnétique. Mathématiquement, cette unification est liée au groupe de symétrie  $SU(2)_L \times U(1)_Y$ . Ces deux groupes  $SU(2)_L$  et  $U(1)_Y$  correspondent respectivement au groupe de jauge d'isospin faible et au groupe de jauge d'hypercharge faible<sup>3</sup>. Le premier groupe est généré par les bosons de jauge  $W_1$ ,  $W_2$  et  $W_3$  alors que le deuxième est généré par le boson de jauge  $B$ .

Physiquement les bosons vecteurs de cette interaction électrofaible sont les bosons  $W^\pm$ , mélanges des bosons  $W_1$  et  $W_2$  et le boson  $Z$  et le photon, obtenus par mélange des bosons de jauge  $W_3$  et  $B$ .

Ce modèle de l'interaction électrofaible décrit complètement et dans un même formalisme les deux interactions faible et électromagnétique. Toutefois cette théorie décrit des bosons  $W^\pm$  et  $Z$  sans masse ce qui est contredit par l'expérience. Or les termes de masse des bosons et des fermions ne sont pas invariants sous la symétrie définie par  $SU(2)_L \times U(1)_Y$ . Il est donc nécessaire de briser cette symétrie.

Le mécanisme qui prévoit cette brisure de symétrie a été théorisé simultanément par trois physiciens qui lui ont donc donné leur nom : Brout, Englert et Higgs. La brisure de symétrie est réalisée par l'introduction d'un champ scalaire dans les équations qui décrivent l'interaction électrofaible. Cet ajout implique l'introduction de masses aux bosons et aux fermions. L'interaction des particules avec ce champ suppose la création d'un nouveau boson : le boson de Higgs.

## L'interaction gravitationnelle

Cette dernière interaction fondamentale est l'interaction qui lie les différentes particules massives entre elles. La particularité de cette interaction est que pour l'instant on ne sait pas s'il existe un boson qui lui est lié (l'existence d'un graviton est purement hypothétique). Elle a une portée infinie mais que son intensité est, relativement à celle des autres interactions fondamentales, beaucoup plus faible si l'on considère des distances de l'ordre de  $10^{-18}$  m. La gravitation est à l'heure actuelle décrite par la théorie de la relativité générale d'Einstein qui n'est pas quantique (ce qui empêche toute unification avec les autres théories).

Le Tableau 1.2 résume les caractéristiques des 4 forces fondamentales.

### 1.1.3 Les limites du Modèle Standard

A l'heure actuelle, le Modèle Standard décrit avec précision et justesse la majorité des phénomènes physiques que nous pouvons observer. Toutefois certains éléments ne sont pas expliqués par ce modèle :

<sup>2</sup>Le  $L$  est présent pour signifier que seuls les fermions gauches (*Left*) et les antifermions droits se couplent avec les courants chargés de l'interaction faible

<sup>3</sup>L'hypercharge faible  $Y$  lie la charge électrique  $Q$  à l'isospin faible  $I_3$  d'une particule :  $Y = 2(Q - I_3)$ .

TAB. 1.2: Interactions fondamentales

Interaction	Bosons porteurs	Charge	Portée (en m)	Intensité ( à $10^{-18}$ m )
Faible	$W^\pm$ / $Z$	Isospin faible	$\leq 10^{-16}$	0.8
Electromagnétique	Photon ( $\gamma$ )	Electrique	$\infty$	1
Forte	Gluons	Couleur	$\leq 10^{-18}$	25
Gravitation	Graviton (hypothétique)	Masse	$\infty$	$10^{-41}$

- Le Modèle Standard ne peut rendre compte de la présence seule de la matière dans l’univers tel que nous le connaissons : l’excédent de matière par rapport à l’antimatière devrait être, selon ce modèle, beaucoup plus faible que celui observé.
- La masse des neutrinos devrait être nulle d’après ce modèle.
- Seules trois des quatre interactions fondamentales sont quantifiées dans ce modèle qui n’inclut pas la gravitation.
- Il n’existe aucune explication convaincante quant à la hiérarchie des masses, le quark *top* étant par exemple un million de fois plus lourd que l’électron ou cent fois plus lourd que le quark *up* alors que ce sont tous des fermions.

Beaucoup de modèles théoriques de Nouvelle Physique sont développés à l’heure actuelle pour tenter de combler les lacunes du Modèle Standard.

## 1.2 Phénoménologie du boson de Higgs

La découverte du boson de Higgs (avec une masse d’environ  $125.9 \text{ GeV}/c^2$  [3]) au LHC en Juillet 2012 ouvre un nouveau champ de recherche pour tester la validité du Modèle Standard et en observer les limites. Ceci peut être effectué par l’étude des différentes caractéristiques de ce boson et en particulier ses couplages avec les autres particules.

### 1.2.1 Constante de couplage du boson de Higgs à un fermion

La constante de couplage d’un fermion avec le boson de Higgs peut être estimée selon le modèle de Yukawa par la formule :

$$g_f = \frac{\sqrt{2}m_f}{\nu} \quad (1.3)$$

où  $m_f$  est la masse du fermion considéré et  $\nu$  la valeur moyenne dans le vide du champ de Higgs <sup>4</sup>. Ainsi plus une particule est lourde, plus sa constante de couplage est importante, cette valeur étant bornée à 1. Le quark *top*, ayant une masse de  $173.07 \text{ GeV}/c^2$  [3], sa constante de couplage avec le boson de Higgs devrait valoir environ 1.

### 1.2.2 Les voies de productions et désintégrations du boson de Higgs

Il est prévu théoriquement que selon la masse du boson de Higgs celui-ci puisse être produit selon différents processus comme indiqué sur la figure 1.1. Il faut aussi prendre en compte le fait que le boson de Higgs ait plusieurs voies de désintégrations possibles décrites sur la Figure 1.2. Chaque mode de

<sup>4</sup>Constante de la théorie de Brout Englert et Higgs dont la valeur est de  $246 \text{ GeV}/c^2$ .

désintégration a une probabilité différente d'être réalisé et on appelle cette probabilité un **rapport de branchement** .

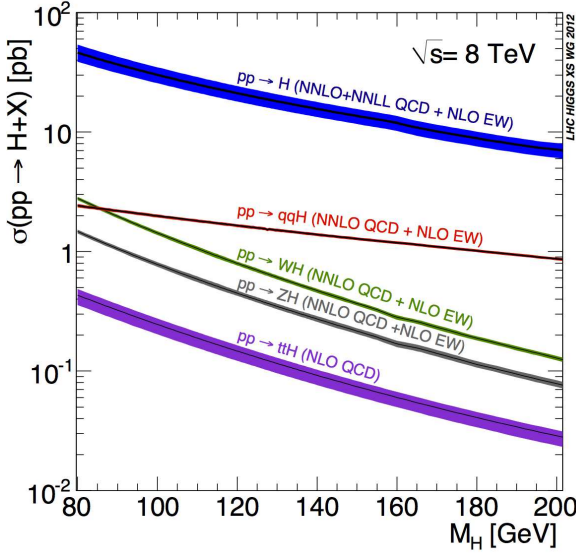


FIG. 1.1: Section efficace des différents modes de production du boson de Higgs

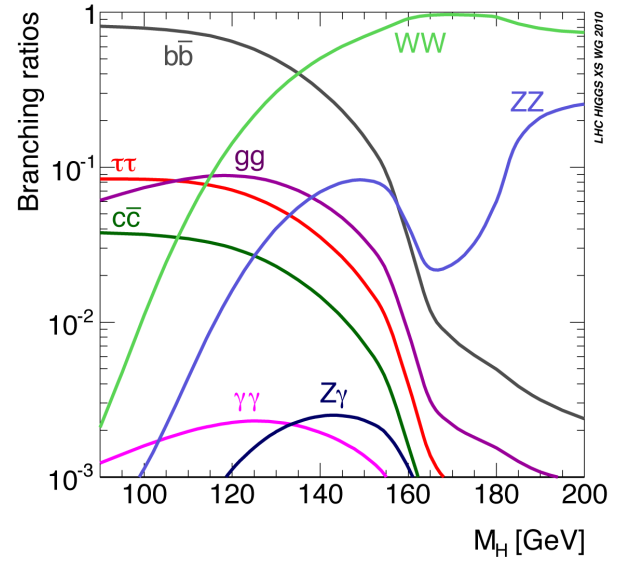


FIG. 1.2: Rapport de branchement du boson de Higgs

### 1.2.3 Le choix de l'étude du processus $t\bar{t}H$

Les différents modes de production du boson de Higgs fournies dans la Figure 1.1 présentent différents couplages du boson de Higgs. Ainsi le couplage du boson de Higgs au quark *top* est présent dans le principal mode de production du boson de Higgs à une masse de 125.9 GeV/c<sup>2</sup>. La production s'effectue par fusion de gluons, en passant par une boucle de quarks *top*. La production du boson de Higgs se fait aussi par couplage avec un boson vecteur (*W* ou *Z*) ou avec une paire de quarks *top*.

La désintégration du boson de Higgs permet aussi la mesure du couplage du boson de Higgs à d'autres particules. Ainsi le couplage du boson de Higgs aux bosons *W* et au boson *Z* ont été observés au LHC par les désintégrations du boson de Higgs, respectivement, en  $W^\pm W^\pm$  et  $ZZ$ . Ensuite, la désintégration du boson de Higgs en deux photons qui a été observée offre la possibilité d'une mesure indirecte du couplage du boson de Higgs au quark *top* par la présence de boucles de quarks *top* dans le diagramme correspondant. Cette dernière voie de désintégration soulève le problème de ce qu'il se passe précisément dans la boucle car deux phénomènes non compris peuvent s'y compenser ce qui ne permet pas une mesure du couplage du quark *top* au boson de Higgs indépendante du modèle.

L'étude du processus  $t\bar{t}H$  offre l'opportunité d'une mesure directe de ce couplage par la mesure de la section efficace<sup>5</sup> de ce processus. Une déviation significative du nombre d'événements observés par rapport au nombre d'événements attendus en considérant les propriétés du boson de Higgs issu du Modèle Standard serait un signe de Nouvelle Physique.

<sup>5</sup>La section efficace est la probabilité qu'un processus physique donné ait lieu et est exprimée en barn (où 1 b=10<sup>-24</sup> cm<sup>2</sup>).

## 2 Outils expérimentaux : Le LHC et le détecteur ATLAS

### 2.1 Le LHC et les collisions proton-proton

#### 2.1.1 L'accélérateur

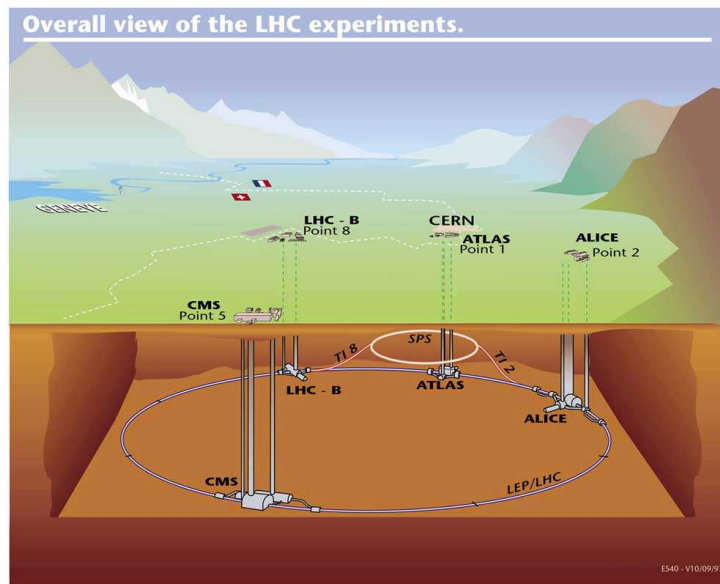


FIG. 2.1: Le LHC et les expériences présentes autour.

Aujourd'hui la physique des particules a comme principal terrain de recherche le LHC (Large Hadron Collider). Cet accélérateur de particules circulaire de 27 km de circonférence est situé au CERN à la frontière franco-suisse (comme indiqué sur la Figure 2.1). Il permet de faire entrer en collision des paquets de protons à une vitesse proche de celle de la lumière et à une énergie dans le référentiel du centre de masse au point de collision de 8 TeV. Actuellement ce collisionneur est à l'arrêt pour le préparer à une seconde phase où l'énergie dans le référentiel du centre de masse devrait atteindre 14 TeV. Cette collaboration unique regroupe plus de 10000 chercheurs et ingénieurs venant de plus de 100 pays différents. Autour de cet anneau sont installés quatre détecteurs différents correspondant chacun à une collaboration en particulier : ATLAS (A Toroidal LHC ApparatuS) et CMS (Compact Muon Solenoid) qui sont deux expériences généralistes, LHCb qui étudie plus particulièrement la physique liée au quark beau et l'asymétrie matière/antimatière et ALICE (A Large Ion Collider Experiment) qui étudie le plasma quark-gluon avec des collisions d'ions lourds (un mois par an est consacré à l'accélération d'ions lourds au LHC). Grâce aux résultats des collisions, les physiciens veulent tester la validité du Modèle Standard et trouver des indices de Nouvelle Physique.

### 2.1.2 Les collisions protons-protons

Le LHC est un collisionneur de protons ce qui implique des propriétés particulières pour les collisions s'y déroulant.

Lors de chaque croisement des paquets de protons il y a le plus souvent plusieurs collisions comme il est possible de le voir sur la Figure 2.2. Ces événements simultanés impliquent un grand nombre de particules produites, laissant des traces dans le détecteur. Ces traces pouvant contaminer la mesure des propriétés des particules qui nous intéressent. Ce phénomène s'appelle l'**empilement** ou le *pile-up*.

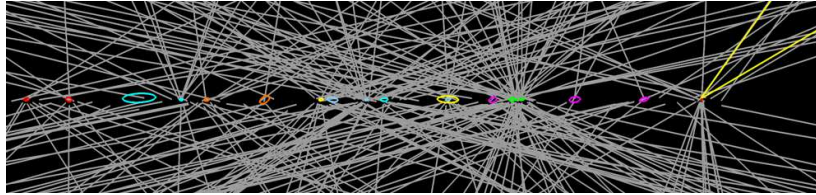


FIG. 2.2: Exemple de *pile-up* observé lors du croisement des deux faisceaux de protons dans le détecteur ATLAS. La reconstruction des traces des particules chargées laisse apparaître plusieurs collisions simultanées.

Lors d'une collision, ce sont les partons (gluons et quarks) des protons qui interagissent. Ainsi les événements que l'on va observer sont les collisions entre deux quarks, entre un quark et un gluon ou entre deux gluons. Toutefois les autres partons des protons peuvent interagir et donner lieu à une collision secondaire ou **événement sous-jacent** (à l'interaction principale) et qui peut être visible dans le détecteur. Lors des collisions entre partons l'impulsion totale est conservée, or nous connaissons l'impulsion totale des protons mais pas celle de chaque parton individuellement. Pour ceux-ci nous savons que l'impulsion totale dans le plan perpendiculaire au faisceau est nulle. Ainsi nous considérons l'impulsion des particules issues de la collision dans ce plan qui sera nommée impulsion transverse et notée  $p_T$ .

La capacité de recherche de processus rare va dépendre du nombre de collisions par seconde et par  $\text{cm}^{-2}$  dans le détecteur. Cette valeur s'appelle la luminosité instantanée. Pour un processus donné le nombre d'événement attendu est égale à :

$$N = \mathcal{L} \times \sigma \quad (2.1)$$

où  $\sigma$  est la section efficace du processus et  $\mathcal{L}$  la luminosité intégrée. Cette dernière est calculée à partir de l'intégrale en temps de la luminosité instantanée et son unité est le barn inverse noté  $\text{b}^{-1}$ . Elle quantifie aussi la taille des données collectées par un détecteur.

## 2.2 ATLAS un détecteur de physique des particules

ATLAS est l'un des quatre détecteurs situés autour du LHC. Il est schématisé sur la figure 2.3. La collaboration qui s'y rapporte regroupe environ 3000 chercheurs du monde entier qui étudient les données fournies par cet appareil titanesque de 45 m de long, 25 m de haut pour un poids total de 7000 tonnes. Pour pouvoir observer les différentes particules qui le traversent, ce détecteur est constitué de plusieurs couches de sous-détecteurs (on compare cette structure à celle d'un oignon), ayant chacune un rôle précis.

### 2.2.1 Le détecteur interne

Ce sous-détecteur est un trajectographe, il a pour but de reconstruire la trajectoire des particules chargées. La présence d'un champ magnétique va permettre de courber la trajectoire de la particule ob-

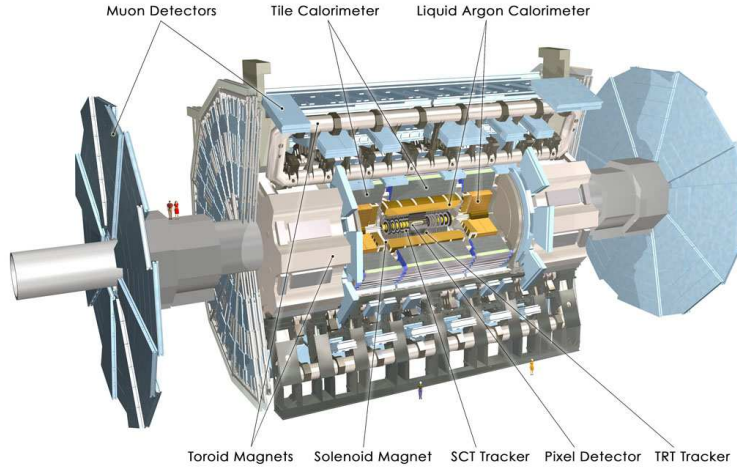


FIG. 2.3: Schéma de détecteur ATLAS.

servée (si cette particule est chargée) et on pourra alors déduire son impulsion et sa charge d'après son rayon de courbure. De plus ce sous-détecteur permet de reconstruire les vertex primaire et secondaires de la collision. Le vertex primaire est l'endroit de l'interaction entre des constituants des deux protons des faisceaux qui ont initialement interagis. Les vertex secondaires sont les différents endroits où des particules ayant des temps de vie assez long pour qu'elles puissent se déplacer dans le détecteur se sont désintégrées. Ce dernier point concerne particulièrement les particules belles et l'existence de ces vertex secondaires joue un rôle important dans le *b-tagging*, c'est-à-dire la reconstruction d'une particule belle dans l'événement.

Ce détecteur interne est constitué lui-même de trois sous-détecteurs différents : un détecteur à pixels, un détecteur à micropistes de silicium (ou SCT) et un trajectographe à rayonnement de transition (ou TRT). Le premier est constitué d'environ 80 millions de pixels disposés sur trois couches cylindriques autour de l'axe du faisceau. La première de ces couches en est située à 5 cm. Ces trois couches devraient permettre d'obtenir trois points pour chaque particule venant du point d'impact.

Le SCT est quant à lui constitué de huit couches de micropistes de silicium. Grâce à celles-ci, huit traces supplémentaires par particule devraient être visibles.

Enfin le TRT est constitué à la fois de brins de polypropylène et de gaz (mélange de Xénon, dioxygène et dioxyde de carbone) et son but est de différencier les traces venant d'électrons et celles venant de pions chargés<sup>1</sup>, cela en étudiant les photons émis par les électrons et les pions lors de leur changement de milieu.

## 2.2.2 Les calorimètres

Les calorimètres ont pour but d'arrêter les particules en dégradant leur énergie par interactions successives avec la matière qui les compose. Par la même occasion, cela va permettre de mesurer leurs énergies. Les dépôts d'énergie vont se faire par création d'une cascade de particules. Les cascades dues à des photons ou des électrons sont qualifiées de cascade électromagnétique. Celles créées par des quarks et des hadrons sont appelées cascades hadroniques. Ces dernières ont aussi une composante électromagnétique.

### Le calorimètre électromagnétique

Le calorimètre électromagnétique est formé de couches superposées de plomb et d'argon liquide. Son but est principalement d'arrêter les particules électromagnétiques que sont les électrons et les photons. En entrant dans ce calorimètre ces particules vont y déposer toute leur énergie. Ce dépôt d'énergie va se faire

<sup>1</sup>Les pions sont des mésons constitués d'un quark  $u$  et d'un antiquark  $d$  ( $\pi^+$ ) ou d'un antiquark  $u$  et d'un quark  $d$  ( $\pi^-$ )

par la création d'une paire électron/positron par un photon, la création d'un photon par collision entre un électron et un positron, l'effet photoélectrique et l'effet Compton. La différenciation entre les électrons et les photons se fait en regardant s'il existe une trace d'une particule chargée dans les trajectographes précédant le calorimètre et qui semble menée à la gerbe électromagnétique que l'on observe. La forme de la gerbe électromagnétique est aussi un critère discriminant pour la différenciation des électrons et des photons dans le calorimètre. Les hadrons et les quarks laisseront aussi des traces dans ce calorimètre mais n'y seront que rarement arrêtés.

## Le calorimètre hadronique

Ce sous-détecteur a pour but d'arrêter toutes les particules hadroniques (telles que les protons ou les pions) produites lors de la collision et de mesurer leurs énergies. Pour cela il est composé de couches superposées de fer et de scintillateurs eux-même reliés à des photomultiplicateurs. Il est beaucoup plus épais que le calorimètre électromagnétique afin de contenir les cascades hadroniques qui vont plus loin dans le détecteur que les cascades électromagnétiques.

### 2.2.3 Le spectromètre à muons

Les seules particules capables de traverser l'ensemble des sous-détecteurs précédents sont les neutrinos et les muons. Or les premiers sont invisibles dans le détecteur car interagissant très rarement avec la matière. Donc les seules particules qui devraient être visibles dans ce spectromètre sont les muons qui interagissent suffisamment avec la matière pour laisser des traces et sortir du détecteur. Le spectromètre à muon permet de reconstruire la trajectoire des muons en plus des détecteurs internes. Il donne donc, en plus de l'identification, une précision supplémentaire sur l'impulsion des muons.

### 2.2.4 Le système de déclenchement

La fréquence de croisement des faisceaux au LHC est de l'ordre de 40 MHz. La quantité de données extraites des collisions qui se déroulent lors de ces croisements est beaucoup trop importante pour que toutes soient stockées. Pour pouvoir faire face à ce problème un système de déclenchement (ou *trigger* en Anglais) a été mis en place. Ce système fonctionne sur deux niveaux. Le premier niveau effectue une sélection grossière des événements qui vont posséder certaines caractéristiques dans les calorimètres et les chambres à muons comme un certain nombre de jets. Ce premier niveau de sélection ramène le nombre d'événements considérés à 75000 par seconde. Ensuite le second niveau du système de déclenchement va s'appuyer sur la reconstruction d'objets tels que les électrons ou les vertex pour diminuer le taux d'événements à 2 kHz. Après ces deux niveaux du système de déclenchement, un troisième tri des événements est réalisé par l'*Event Filter* pour arriver au final au stockage de 200 événements par seconde.

### 2.2.5 Conventions géométriques

Des conventions sont utilisées pour décrire les quantités géométriques dans le détecteur. L'axe  $z$  est placé suivant l'axe du faisceau, les coordonnées  $x$  et  $y$  sont décidées de telle façon à avoir un repère direct. L'angle  $\phi$  est défini dans le plan dit transverse ( $O, x, y$ ) comme étant l'angle par rapport à l'axe  $y$ . L'angle  $\theta$  est l'angle par rapport à l'axe  $z$ . Toutes ces grandeurs sont indiquées sur la Figure 2.4. La pseudorapidité  $\eta$  est préférée à l'angle  $\theta$ . Cette grandeur est calculée d'après l'équation suivante :

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right) \quad (2.2)$$



En prenant  $\theta$  entre  $\pi/16$  et  $\pi/2$  on a  $\eta$  qui va évoluer de -2.5 à 0.  
A partir de la définition de ces deux angles,  $\eta$  et  $\phi$ , il est possible de définir la distance entre deux particules comme étant :

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \quad (2.3)$$

où  $\Delta\eta$  est la différence des pseudorapidités et  $\Delta\phi$  la différence des angles  $\phi$ .

On utilisera dans la suite la notion de grandeurs transverses, telle que l'impulsion transverse qui est la projection du vecteur impulsion dans le plan transverse.

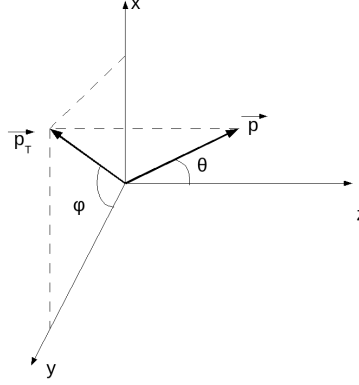


FIG. 2.4: Système de coordonnées utilisé dans ATLAS. Le vecteur  $\vec{p}$  représente l'impulsion de la particule alors que le vecteur  $\vec{p}_T$  représente son impulsion transverse.

## 2.2.6 ATLAS, une collaboration internationale et une équipe à Clermont-Ferrand

La collaboration ATLAS regroupe à l'heure actuelle quelques 3000 chercheurs venus de 174 instituts différents, basés dans 38 différents pays. L'équipe ATLAS présente au LPC à Clermont compte pour sa part 20 membres sous la direction de Dominique Pallin, directeur de recherche au CNRS. Les membres de cette équipe remplissent plusieurs tâches auprès d'ATLAS telles que le traitement et l'analyse des données fournies par le LHC ou l'entretien et l'amélioration du calorimètre hadronique.

Le LPC de Clermont-Ferrand, structure d'accueil de l'équipe ATLAS, est une Unité Mixte de Recherche de L'Université Blaise Pascal et du CNRS et est affilié à l'IN2P3 (Institut National de Physique Nucléaire et de Physique des particules). Son effectif comprend 21 chercheurs et 33 enseignants-chercheurs ainsi que des ingénieurs et du personnel administratif. Il faut ajouter à cela une population importante de doctorants et post-doctorants. Le laboratoire est impliqué dans 14 expériences différentes. Sur ce total trois équipes sont concernées par les recherches menées au LHC : les équipes ALICE, LHCb et ATLAS.

## 2.3 Les données utilisées dans cette analyse

Les données utilisées sont les données collectées par le détecteur ATLAS lors des collisions proton-proton sur l'ensemble de l'année 2012 à une énergie dans le référentiel du centre de masse de 8 TeV ce qui correspond à une luminosité intégrée de  $20.3 \text{ fb}^{-1}$ .

# 3 Signature choisie et préselection pour l'étude du processus $t\bar{t}H$

Le processus étudié est la production d'un boson de Higgs associé à une paire de quark *top*. Par la suite nous parlerons de signal dans le cas du processus  $t\bar{t}H$  et tout autre processus sera un bruit de fond. Le but de l'analyse est de trouver une région où l'écart entre le nombre d'événements de signal et de bruit de fond sera suffisant pour les distinguer. Pour cela, les caractéristiques des bruits de fond et du signal doivent être bien connues. Une simulation Monte-Carlo de ces processus est réalisée pour répondre à cette attente. Après une préselection grossière, la région de signal est définie à l'aide de sélections sur des variables discriminantes.

## 3.1 Les objets reconstruits dans ATLAS

Pour pouvoir reconnaître un processus lors d'une désintégration dans le détecteur ATLAS certains objets doivent être reconstruits :

- Les **électrons** (et positrons) sont reconstruits à partir d'une gerbe dans le calorimètre électromagnétique, d'une trace chargée avant le calorimètre et de la forme de la distribution de l'énergie déposée par la particule dans le calorimètre.
- Les **muons** (et antimuons) sont déduits d'une trace présente dans les chambres à muons et d'une trace dans le détecteur interne.
- L'**énergie transverse manquante** (*Missing Energy Transverse* ou MET en anglais) est la somme des impulsions transverses déposées par les différentes particules. La somme des impulsions est nulle dans le plan transverse avant la collision, elle doit donc toujours l'être après par conservation de l'énergie. Ainsi, si cette grandeur est non nulle ceci est la preuve qu'une particule invisible a été produite (généralement des neutrinos).
- Les **jets** sont dus à la fragmentation puis à l'hadronisation de quarks produits par la collision. Ils vont laisser des gerbes dans le calorimètre hadronique.
- Les jets étiquetés *b* ou *bjets* sont des jets dus à la désintégration d'une particule belle (qui contient un quark *b*). Pour reconnaître cet objet on regarde s'il existe un vertex secondaire et on regarde si sa masse est relativement importante (les particules belles ayant une masse supérieure à 5 GeV/c<sup>2</sup>).

Les critères considérés pour la reconstruction des électrons, des muons et des jets sont différents et varient d'une analyse à l'autre. Dans notre cas, les caractéristiques d'un objet pour qu'il soit reconstruit comme un électron sont :

- une énergie transverse supérieure à 25 GeV : réduit la contribution d'électron mal reconstruit ou qui n'en sont pas et correspond au seuil du *trigger* (24 GeV).
- une pseudorapidité  $|\eta| < 2.47$  et  $\eta \notin ]1.37; 1.52[$  : zone moins instrumentée d'ATLAS (pour les calorimètres).
- le respect des caractéristiques d'un électron *tight* comme défini dans le groupe ATLAS [4] : rejets des

électrons faux (reconstruits à partir de jets).

- une distance inférieure à 2mm du vertex primaire en  $z$  : contre l'effet du pile-up.
- le respect de la *mini-isolation* : la somme des impulsions transverses des traces (excepté celle de l'électron) dans un rayon  $\Delta R = \min(0.4; 10 \text{ GeV}/p_T)$  autour de la trace de l'électron doit être inférieure à 0.05 fois l'impulsion transverse de l'électron.
- une séparation entre l'électron et le jet le plus proche de  $\Delta R = 0.4$  : exclut (avec la *mini-isolation*) les électrons provenant des jets.

Cette sélection définit des électrons que nous qualifierons par la suite de *tight*. Nous utiliserons aussi des électrons dits *loose*, ce qui correspond à la même sélection que pour les électrons *tight* sans la mini-isolation et en utilisant la définition des électrons *medium* dans ATLAS [4].

Les pré-requis pour la reconstruction d'un objet comme étant un muon *tight* sont :

- une impulsion transverse supérieure à 25 GeV : réduit la contribution des muons mal reconstruit ou qui n'en sont pas et correspond au seuil du *trigger* (24 GeV).
- une trajectoire dans la région  $|\eta| < 2.5$  : zone moins instrumentée d'ATLAS.
- le respect du même critère de mini-isolation que les électrons.
- une séparation entre le muon reconstruit et le jet le plus proche de  $\Delta R > 0.04 + 10 \text{ GeV}/p_T$  : exclut (avec la *mini-isolation*) les muons provenant des jets
- un rapport entre le paramètre d'impact par rapport à l'axe du faisceau dans le plan transverse, (d0) et son incertitude  $< 3$  : pour limiter le nombre de muons non prompt.

De même que pour les électrons, on peut définir des muons *loose* en supprimant le critère de mini-isolation. Les jets sont très présents dans le détecteur ATLAS à cause du phénomène de *pile-up* et du fait que le LHC soit un collisionneur proton-proton. Pour diminuer la contribution des jets venant de cet empilement on définit la grandeur *jet-vertex-fraction* ou JVF. C'est le rapport entre la somme de l'impulsion transverse (si celle-ci est supérieure à 0.4 GeV) de toutes les traces qui originaires du vertex primaire qui composent le jet et la somme des impulsions transverses de toutes les traces qui sont contenues dans le jet. Les objets reconstruits comme des jets dans cette analyse ont comme caractéristiques :

- une impulsion transverse supérieure à 25 GeV.
- une pseudorapidité  $|\eta| < 2.5$  : zone moins instrumenté d'ATLAS.
- $JVF > 0.5$  si l'impulsion du jet est inférieure à 50 GeV et  $|\eta| < 2.4$  : contre l'effet du *pile-up*.
- une séparation entre le jet reconstruit et l'électron *loose* le plus proche de  $\Delta R > 0.2$  : réduit la reconstruction de jets comme étant des électrons.

### 3.2 Choix de la signature expérimentale : deux leptons de même signe

Dans le cadre de cette étude nous regarderons les cas où l'état final contient deux leptons de même signe. Ce choix est motivé par la faible quantité de processus physiques autres que le processus  $t\bar{t}H$  qui peuvent donner un tel état final dans le Modèle Standard.

Les leptons qui nous intéressent ici sont les électrons et les muons car les leptons  $\tau$  se désintègrent avant d'être visibles dans le détecteur et cela principalement sous forme de hadrons. Ainsi les trois canaux qui seront envisagés sont les canaux  $e^\pm e^\pm$ ,  $e^\pm \mu^\pm$  et  $\mu^\pm \mu^\pm$ .

Les voies de désintégration du boson de Higgs qui vont nous donner un leptons sont  $H \rightarrow W^+ W^-$ ,  $H \rightarrow ZZ$  et  $H \rightarrow \tau^\pm \tau^\pm$ . Un seul lepton est demandé venant du Higgs pour avoir deux leptons de même signe car nous devons en avoir un qui vienne d'un boson W issu d'un quark *top* (voir section 1.1.2).

Le Tableau 3.1 regroupe les probabilités (noté  $P_{ss}$ ) d'avoir deux leptons de même signe dans l'état final du processus  $t\bar{t}H$  et le rapport de branchement [5] du boson de Higgs pour les trois canaux de désintégration considérés.

Le nombre d'événement du processus  $t\bar{t}H$  qui comporte deux leptons de même signe dans l'état final

TAB. 3.1: Probabilité pour le processus  $t\bar{t}H$  de se désintégrer en deux leptons de même signe et rapport de branchement du Higgs dans les voies de désintégrations du boson de Higgs en  $W^\pm W^\pm, \tau^\pm \tau^\pm$  et  $ZZ$

	$H \rightarrow WW$	$H \rightarrow \tau\tau$	$H \rightarrow ZZ$
$P_{ss}$	0.0713	0.0861	0.0054
BR	0.228	0.062	0.029

pour une voie de désintégration donnée est le produit de la probabilité  $P_{ss}$  par le rapport de branchement du boson de Higgs dans la voie de désintégration voulue. Au final le pourcentage d'événements avec deux leptons de même signes pour le processus  $t\bar{t}H$  selon la voie de désintégratoin est présentée dans le Tableau 3.2.

TAB. 3.2: Proportion (en %) d'événement avec deux leptons de même signe dans l'état final  $t\bar{t}H$  selon les voies de désintégration du boson de Higgs en  $W^\pm W^\pm, \tau^\pm \tau^\pm$  et  $ZZ$

$H \rightarrow WW$	$H \rightarrow \tau\tau$	$H \rightarrow ZZ$
75%	24%	1%

La voie de désintégration dominante du boson de Higgs pour cette étude est donc la voie en  $W^+W^-$ . Dans la suite de l'étude nous qualifierons d'« hadroniques » les bosons  $W^\pm$  qui se désintègrent en deux quarks dans le détecteur et par extension, ce qualificatif sera appliqué au quark *top* se désintégrant sans lepton dans l'état final. Les bosons  $W^\pm$  qui donneront un lepton et un neutrino seront dits « leptoniques » et les quarks *top* correspondant « semi-leptoniques ».

Finalement nous attendrons dans l'état final deux leptons de même signe, deux neutrinos (invisibles dans le détecteur) et six jets dont deux étiquetés *b* comme indiqué sur le diagramme de la Figure 3.1.

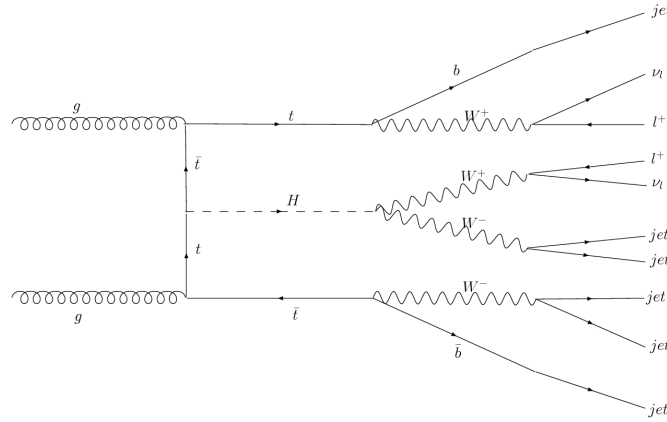


FIG. 3.1: Diagramme de production  $t\bar{t}H$  dont la désintégration aboutit à la production de deux leptons de même signe avec  $g$  pour gluon,  $t$  pour quark *top*,  $b$  pour quark beau,  $l$  pour lepton et  $H$  pour boson de Higgs.

### 3.3 Génération des événements Monte-Carlo

Pour notre analyse nous devons modéliser le signal et les bruits de fonds. Pour cela des événements Monte-Carlo sont générés avec différents générateurs tels que MADGRAPH 5.1[6], PYTHIA[7] ou SHERPA[8] (pour la partie calcul, fragmentation, hadronisation et désintégration) et GEANT4[9] pour la simulation du détecteur.

Ces échantillons sont générés à des luminosités intégrées supérieures à celle des données et le nombre d'événements que nous aurons pour les différents processus devra être renormalisé pour une luminosité intégrée de  $20.3 \text{ fb}^{-1}$ , luminosité intégrée délivrée à ATLAS par le LHC en 2012. Un poids de normalisation sera attribué à chaque échantillon généré. Ce poids sera calculé par le rapport entre la luminosité intégrée équivalente et celle obtenue par les données.

### 3.4 Bruit de fond

Le choix d'une signature avec deux leptons de même signe dans l'état final permet de limiter le nombre d'événements de bruit de fond que l'on pourrait avoir. Toutefois quelques processus peuvent mimer cette signature.

Ces bruits de fond peuvent se diviser en deux catégories. Tout d'abord les bruits de fond irréductibles ayant la même signature que le signal. La seconde catégorie est l'ensemble des processus qui n'ont pas deux leptons de même signe dans l'état final mais qui sont reconstruits comme si c'était le cas. Cette dernière catégorie peut se diviser en deux processus différents : les événements Mis-Id et les événements Fakes.

#### 3.4.1 Les incertitudes sur les bruits de fond simulés par Monte-Carlo

Les incertitudes sur le nombre d'événements générés par Monte-carlo peuvent se grouper en deux catégories :

- Une **incertitude statistique** liées à la taille limitée de ces échantillons.
- Une **incertitude systématique** liée à la normalisation à la luminosité de la section efficace des processus générés<sup>1</sup>.

Le Tableau 3.3 présente ces différentes incertitudes systématiques pour le bruit de fond généré par Monte-Carlo.

TAB. 3.3: Incertitudes systématiques (en %) sur le nombre d'événement des bruits de fonds

$t\bar{t}V + jets$	$t\bar{t}W$	$WW$	$WZ$	$ZZ$	$V (W \text{ ou } Z) + jets$	$t\bar{t}$
$\pm 30\%$	$\pm 40\%$	$\pm 25\%$	$\pm 30\%$	$\pm 30\%$	$\pm 30\%$	$\pm 10\%$

#### 3.4.2 Bruits de fond irréductibles

Ces bruits de fond sont des processus qui vont donner une signature similaire au signal, deux leptons de même signe :

- Triplet  $VVV$  (avec  $V$  un boson  $W^\pm$  ou  $Z$ ).
- Paire  $W^\pm W^\pm + jets$ .
- Paire  $W^\pm Z + jets$ .
- Paire  $ZZ + jets$ .
- Processus  $t\bar{t}W$ ,  $t\bar{t}Z$  et  $t\bar{t}WW$ .
- Couple  $W^\pm H/ZH$ .

Pour un faible nombre de jets les processus  $VV + jets$  seront majoritaires alors que ce seront les processus  $t\bar{t} + V$  pour un grand nombre de jets.

<sup>1</sup>La luminosité intégrée ou la reconstruction des objets sont aussi source d'incertitudes systématiques mais celles-ci sont inférieures à 10% et ne seront pas prises en compte.

### 3.4.3 Les événements *Fakes*

Les événements **Fakes** sont des événements pour lesquels des jets (par exemple si un jet dépose toute son énergie dans le calorimètre électromagnétique) ou des événements pour lesquels les leptons choisis sont des leptons provenant de désintégrations d'autres particules que celles qui nous intéressent (par exemple de particules belles) sont reconstruits comme des leptons venant du vertex primaire. Le principal processus à l'origine de ces événements est la création d'une paire  $t\bar{t}$  dans le détecteur (ainsi que les processus avec un boson  $Z$  ou  $W$  et des jets, principalement dans les régions avec un faible nombre de jets). Ces processus sont difficilement modélisables et doivent être obtenus à partir des données. Toutefois des événements fakes simulés seront dans un premier temps utilisés.

### 3.4.4 Les événements *Mis-Id*

Les événements **MisId** sont des événements qui correspondent à des cas où la charge d'un lepton est mal mesurée :

- lorsque les leptons ont de grandes impulsions donc une trajectoire très peu courbée (la courbure de la trajectoire étant l'un des éléments principaux permettant la reconstruction de la charge d'une particule).
- due à un trident ( $e^- \rightarrow e^- \gamma \rightarrow e^- e^+ e^-$ ) vu comme un électron avec une charge mal attribuée.

Cette catégorie ne regroupe que des événements avec des électrons car la probabilité d'une erreur sur la charge d'un muon est négligeable [10]. La création de paires  $t\bar{t}$  est le processus dominant à grand nombre de jets et les processus  $Z$ +jets domineront à bas nombre de jets. Ces processus se déroulent fortement pour des électrons avec une valeur absolue de la pseudorapidité élevée, au dessus de 1.5 (voir figure 3.2).

Tout comme pour les événements Fakes, de nombreux processus peuvent fournir des électrons dont la charge est mal identifiée et ainsi ce phénomène est difficile à modéliser par simulation Monte-Carlo. Il est donc plus précis d'extraire le nombre d'événements Mis-Id à partir des données. La méthode pour cela consiste à regarder le nombre d'événements avec des électrons de même signe ayant une masse invariante proche de la masse du boson  $Z$  ( $Z \rightarrow e^+ e^-$ ,  $m_Z = 91 \text{ GeV}/c^2$ , voir figure 3.2) par rapport au nombre d'événements que l'on obtient pour des électrons de charges opposées. Le rapport des intégrales des deux pics de masse est utilisé pour calculer la probabilité à ce que la charge d'un électron ne soit pas la bonne. L'incertitude systématique liée à cette méthode est de 30%.

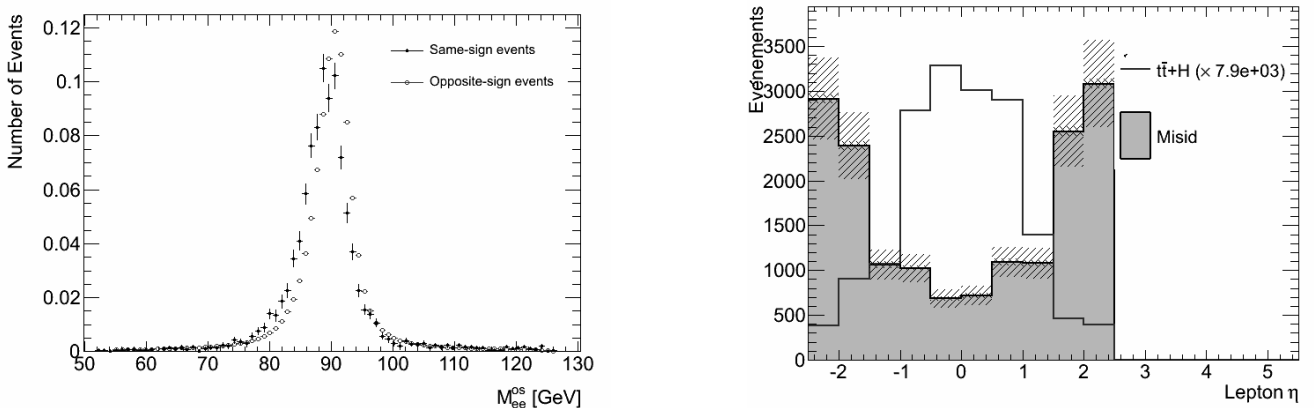


FIG. 3.2: Distribution de la masse invariante et en pseudorapidité des électrons des événements Mis-Id

Nous utiliserons également dans la suite de l'étude une simulation Monte-Carlo des événements Mis-Id.

### 3.5 Modélisation du signal et reconstruction des objets hadroniques

Nous regardons la simulation Monte-Carlo du signal pour estimer notre capacité à reconstruire le quadri-moment des particules de notre état final.

Ceci ne concerne que les particules qui vont donner des quarks en se désintégrant, donc un quark *top* (un boson *W* et un quark *b*) et un boson *W* qui vient du Higgs. Le quark *top* et le boson *W* restants ne sont pas reconstituables à cause de la présence des neutrinos dans l'état final. Malgré la présence dans l'état final de tout les quarks issus des objets hadroniques, leur reconstruction n'est pas assurée par le fait que les quarks forment des jets dans le détecteur qui peuvent être mal reconstruits.

Dans l'échantillon que nous regardons nous avons les informations sur les particules générées ainsi que sur les objets reconstruits, la véracité de la reconstruction d'un jet peut donc être vérifié en choisissant, par exemple, de considérer comme bien reconstruit un jet qui se trouve dans un rayon proche de la particule vraie ( $\Delta R < 0.4$ ).

Cet étiquetage des jets rend compte de notre capacité à reconstruire tous les éléments hadroniques d'un événement. Si on peut reconstruire ces particules, on pourra alors utiliser les quantités reconstruites (telle que la masse invariante de ces particules) comme variables discriminantes.

Pour quantifier cette capacité nous avons considéré la fraction entre le nombre d'événements avec deux leptons simulés et le nombre de ces événements dont tous les quarks peuvent être effectivement associés à des jets observés dans le détecteur.

Après traitement nous observons que ce rapport n'est que de 10% ce qui compromet la reconstruction des éléments hadroniques des événements  $t\bar{t}H$ . Dans les autres cas au moins un des quarks n'a pas été reconstruit (souvent des quarks avec une faible impulsion transverse et ne passant pas les critères précisés dans la section 3.1).

### 3.6 Définition de la signification statistique

En physique des particules les données sont traitées statistiquement et nous allons, dans la suite de cette étude, chercher à maximiser la grandeur statistique qu'est la signification statistique.

La distribution des événements sur un intervalle pour une variable donnée, telle que le nombre de jets pour un nombre d'événements suffisant, va obéir à une distribution gaussienne. L'écart type, noté  $\sigma$ , quantifie la dispersion d'un ensemble de données par rapport à sa moyenne. Connaissant cette valeur nous pouvons ainsi exprimer la probabilité qu'un événement se trouvant à  $n\sigma$  de la moyenne soit un événement de la distribution (comme illustré sur la Figure 3.3).

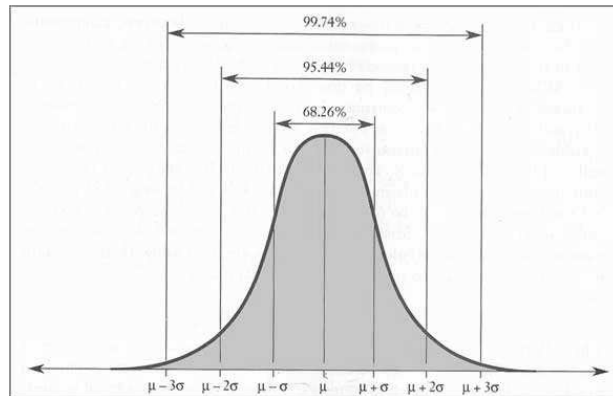


FIG. 3.3: Distribution gaussienne avec le pourcentage d'événements compris dans un intervalle de  $n\sigma$  autour de la valeur moyenne

Pour un processus physique donné, nous pouvons prendre le bruit de fond comme étant la distribution

et le signal plus le bruit de fond (ou les données) comme étant l'événement dont nous voulons connaître la probabilité d'être dans cette distribution. En lien avec ce choix, une convention prise est que pour annoncer la découverte d'une nouvelle particule il faut que l'événement soit à  $5\sigma$  (il y aurait une chance sur plus de 3,5 millions que ce soit une fluctuation statistique du bruit de fond).

Dans le cas d'une distribution gaussienne la signification statistique est  $\sigma_s = \frac{s}{\sqrt{b}}$  avec  $s$  et  $b$  qui sont respectivement les nombres d'événements de signal et de bruit de fond attendus après sélection.

Une analyse (une sélection d'événement) sera d'autant plus capable de révéler un signal superposé au bruit de fond que la signification statistique est grande. On cherchera donc à maximiser cette grandeur. Pour une plus grande précision de la signification statistique, il est possible de la définir à partir d'une loi de Poisson.

### 3.7 Préselection

Notre but est d'obtenir un ensemble d'événements pour lesquels l'état final comporte deux leptons de même signe. Pour cela nous allons effectuer un lot de sélections qui vont constituer une préselection.

Ces sélections sont :

**Trigger :** pour les événements venant des données du LHC c'est le choix du critère ayant déclenché l'enregistrement de l'événement alors que dans le cas de la simulation Monte-Carlo c'est la simulation de ce système de déclenchement.

**2 leptons de même signe et leptons *tight* :** les événements ne sont gardés que si il y a deux leptons et qu'ils obéissent aux critères *tight* définis précédemment

**Correspondance au *trigger* et lepton *tight* :** pour que l'un des leptons de la paire de même signe soit celui qui ait déclenché le trigger (il a donc un  $p_T$  suffisamment grand) et qu'il soit *tight*.

**Sélection en  $|\eta|$  sur les électrons :** les électrons ne sont conservés que si ils sont dans la région  $|\eta| < 1.5$  : limite le nombre d'événements Mis-Id.

**Au moins 1 jet**

Le nombre d'événements des processus irréductibles ainsi que les Fakes et la Mis-Id sont estimés sur simulation et donnés dans les Tableaux 3.4 et 3.5 respectivement.

TAB. 3.4: Nombre d'événements et incertitude statistique pour les bruits de fond irréductibles et le signal après la préselection

	$WH/ZH$	$VVV$	$t\bar{t}V(V) + jets$	$WW$	$WZ$	$ZZ$	$t\bar{t}H$
$e^\pm e^\pm$	$2.27 \pm 0.35$	$0.50 \pm 0.04$	$10.57 \pm 0.27$	$10.96 \pm 0.23$	$80.41 \pm 1.60$	$17.73 \pm 1.45$	$1.60 \pm 0.08$
$e^\pm \mu^\pm$	$7.56 \pm 0.74$	$1.8 \pm 0.07$	$32.65 \pm 0.48$	$39.06 \pm 0.43$	$236.83 \pm 2.80$	$34.93 \pm 2.01$	$4.45 \pm 0.14$
$\mu^\pm \mu^\pm$	$7.77 \pm 0.86$	$1.63 \pm 0.07$	$23.32 \pm 0.42$	$30.72 \pm 0.39$	$150.12 \pm 2.27$	$18.15 \pm 1.47$	$2.68 \pm 0.11$

TAB. 3.5: Nombre d'événements et incertitude statistique pour les bruits de fond Fakes et Mis-Id après la préselection

	<i>Fakes</i>	<i>Mis - Id</i>
$e^\pm e^\pm$	$238.0 \pm 11.5$	$2515.0 \pm 50.7$
$e^\pm \mu^\pm$	$386.65 \pm 15.66$	$107.38 \pm 5.75$
$\mu^\pm \mu^\pm$	$86.45 \pm 5.80$	—



## 4 Estimation et contrôle des bruits de fond

Après avoir effectué la préselection, l'adéquation entre les bruit de fond simulés et les données est testée. Pour tester la validité de cette simulation il existe plusieurs méthodes.

### 4.1 Bruit de fond global

Nous allons regarder si la somme des bruits de fond est compatible avec les données sans distinguer un processus en particulier. Nous pouvons tester ces bruits de fond en comparant notre simulation avec les données. Cette méthode ne peut être utilisée que dans des régions où le nombre d'événements de signal attendu est faible afin de ne pas introduire de biais. Il nous faut donc définir des régions de contrôle avec peu d'événements de signal attendu. La région que nous allons considérer est une région avec moins de 4 jets et aucun jet étiqueté  $b$ . Nous avons aussi relâché la sélection en  $\eta$  sur les électrons tout en excluant les événements dont la masse invariante des deux électrons est située sur un intervalle de 10 GeV autour de la masse du boson  $Z$  et en dessous de 15 GeV. Cette dernière sélection limite la présence du processus  $Z$ +jets qui est très largement dominant dans cette région pour le canal  $e^\pm e^\pm$ .

TAB. 4.1: Nombre d'événements attendu dans la région de contrôle (pour le bruit de fond et le signal)

	Signal(MC)	Mis-Id(MC)	Fakes(MC)	Autres bdf(MC)
$e^\pm e^\pm$	$0.165 \pm 0.028$	$2432.29 \pm 62.52$	$382.83 \pm 15.80$	$173.18 \pm 3.92$
$e^\pm \mu^\pm$	$0.333 \pm 0.039$	$107.88 \pm 6.27$	$552.17 \pm 20.15$	$406.38 \pm 4.21$
$\mu^\pm \mu^\pm$	$0.156 \pm 0.021$	0	$64.25 \pm 5.00$	$189.23 \pm 2.74$

La contamination en signal dans cette région de contrôle étant suffisamment faible (comme noté dans le Tableau 4.1), nous devrions donc avoir un nombre d'événements venant des données sensiblement égal à celui du bruit de fond simulé si celui-ci est bien modélisé. Nous comparons aussi les distributions de différentes variables pour les événements de données et de la simulation. L'écart entre les deux est quantifié en terme de signification statistique. Si le bruit de fond est correctement modélisé, la signification statistique devrait être relativement faible (typiquement  $< 3\sigma$ ).

TAB. 4.2: Nombre d'événements attendus et nombres d'événements dans les données

	Simulation (Fakes/MisId MC)	Données
$e^\pm e^\pm$	$2988.3 \pm 64.6 \pm 344.0$	2598
$e^\pm \mu^\pm$	$1066.4 \pm 21.5 \pm 102.5$	1098
$\mu^\pm \mu^\pm$	$253.5 \pm 5.7 \pm 41.0$	251

Le nombre d'événements dans les données (indiqué dans le Tableau 4.2) est compatible avec notre

modélisation du bruit de fond. Nous considérons les distributions des grandeurs suivantes :

- le nombre de jets
- le nombre de jets étiquetés  $b$
- l'impulsion transverse des leptons
- la pseudorapidité  $\eta$  des leptons

Les distributions de  $H_T$  dans les différents canaux pour les régions de contrôle sont présentées dans la figure 4.1. Les distributions pour les autres variables des événements de bruit de fond et des données sont présentées dans l'Annexe A.

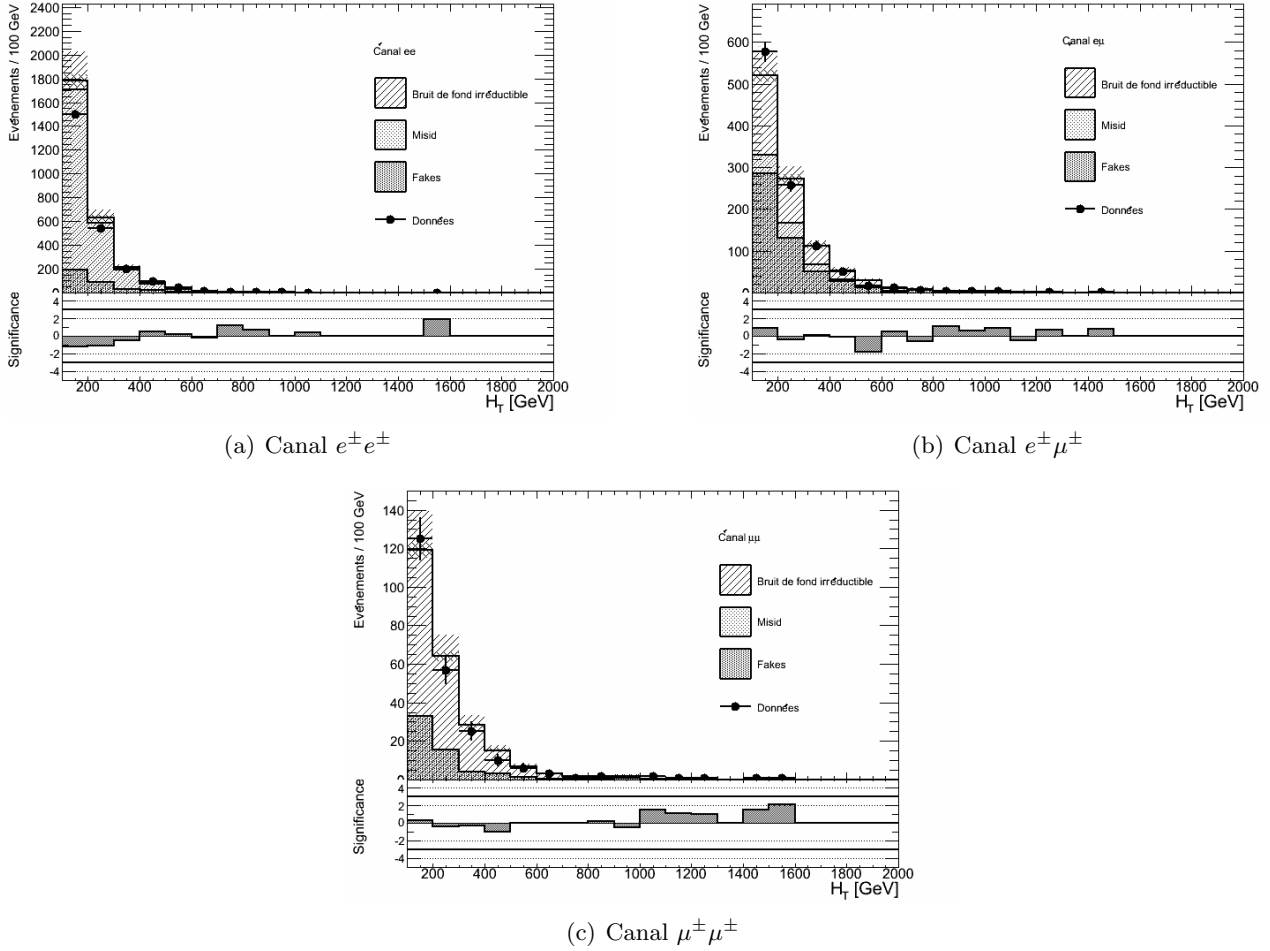


FIG. 4.1: Distributions de la somme des impulsions transverses des leptons et des jets dans les trois canaux pour le bruit de fond et les données dans la région de contrôle avec moins de 4 jets et aucun jets étiquetés  $b$ .

Nous avons dans la région de contrôle un bon accord entre notre simulation des bruits de fond et les données, à la fois pour les distributions de différentes grandeurs et aussi sur le nombre total d'événements. Toutefois ces résultats montrent simplement que le bruit de fond total est correct mais ne présage pas de la modélisation individuelle de chaque bruit de fond, dont les biais peuvent se compenser pour donner un accord de la somme des bruits de fond modélisés avec les données. Nous devons donc définir des régions de contrôles de telle façon que l'on maximise la quantité d'événements due à un bruit de fond particulier par rapport aux autres bruits de fond et au signal ensuite une comparaison aux données permettra d'évaluer la qualité de la simulation du processus choisi.

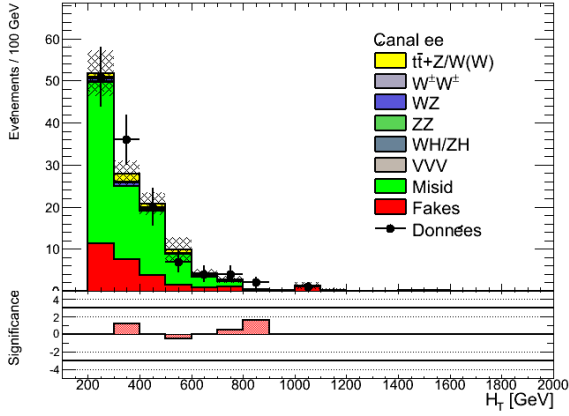
## 4.2 Processus $t\bar{t}V$

Le processus  $t\bar{t}V$  est pour notre analyse un bruit de fond irréductible. La viabilité de la simulation peut être vérifiée en choisissant une région de contrôle dédiée à ce processus. La présence de ce bruit de fond doit y être maximiser tout en limitant la sélection (le but est de savoir si ce processus est bien reconstruit dans une région ec un maximum d'événements) et en limitant le nombre d'événements de signal.

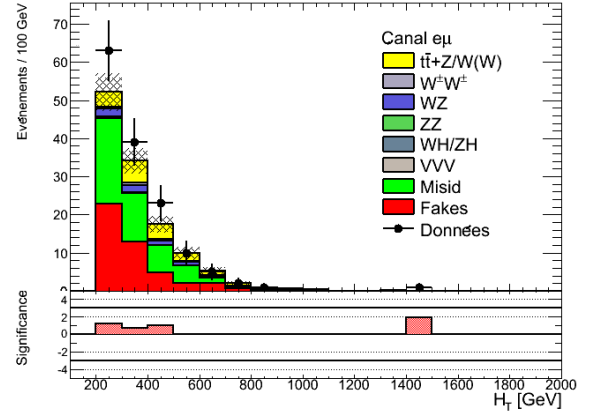
La définition des régions de contrôle qui seront considérées pour ce test varie selon le canal et ces définitions sont regroupées dans le Tableau 4.3.

TAB. 4.3: Sélections effectuées pour la définition des régions de contrôle du  $t\bar{t}V$

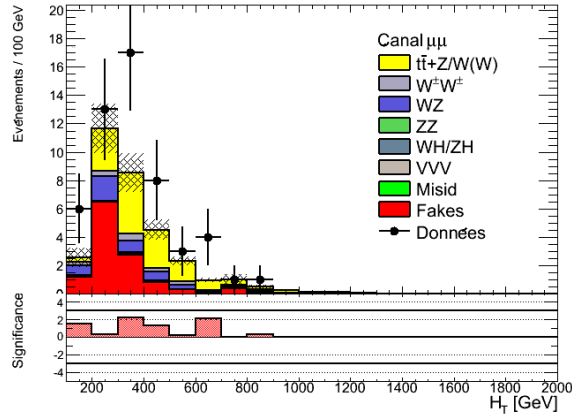
	$ee$	$e\mu$	$\mu\mu$
$N_{jets}$	$\in [2; 4]$	$\in [2; 4]$	$\in [2; 4]$
$N_{bjets}$	$\geq 1$	$\geq 1$	$\geq 1$
$Ht$	$> 200 \text{ GeV}$	$> 200 \text{ GeV}$	
$\eta_{elec}$	$\in ] - 1.5; 1.5[$	$\in ] - 1.5; 1.5[$	



(a) Canal  $e^{\pm}e^{\pm}$



(b) Canal  $e^{\pm}\mu^{\pm}$



(c) Canal  $\mu^{\pm}\mu^{\pm}$

FIG. 4.2: Distributions de la somme des impulsions transverses des leptons et des jets dans les trois canaux pour le bruit de fond et les données dans la région de contrôle pour le processus de bruit de fond  $t\bar{t}V$ .

Les nombres d'événements de bruit de fond et de données dans ces régions de contrôle sont données dans le Tableau 4.4.

TAB. 4.4: Nombre d'événement pour le processus  $t\bar{t}V$ , le bruit de fond total et les données dans les régions de contrôle du  $t\bar{t}V$

	$e^\pm e^\pm$	$e^\pm \mu^\pm$	$\mu^\pm \mu^\pm$
$t\bar{t}V$	$6.30 \pm 0.21 \pm 1.06$	$18.35 \pm 0.37 \pm 3.16$	$13.43 \pm 0.33 \pm 2.43$
$Bdf(t\bar{t}V \text{ compris})$	$119.63 \pm 6.57 \pm 6.31$	$122.97 \pm 5.48 \pm 7.14$	$32.75 \pm 2.03 \pm 2.88$
<i>Données</i>	125	144	53

L'estimation des bruits de fond est en accord avec les données dans les différentes régions et canaux et nous pouvons donc conclure à la bonne modélisation du processus  $t\bar{t}V$ . La Figure 4.2 regroupe les distributions en  $H_T$  pour les trois canaux des bruits de fond et des données. Cette méthode peut être utilisée pour tester la modélisation de chaque bruit de fond intervenant dans l'analyse mais par la suite nous nous sommes penchés sur l'estimation du nombre d'événements Fakes dans la région de signal, définie dans la section suivante.

### 4.3 Estimation du nombre d'événements *Fakes* depuis les données

Le nombre d'événements Fakes est complexe à estimer car venant de nombreux processus différents (paire  $t\bar{t}$ ,  $V$ +jets ...). Il existe plusieurs méthodes pour résoudre ce problème et celle qui va être étudiée dans la suite consiste en l'estimation du nombre d'événements dû à ce processus dans la région de signal à partir du nombre d'événements Fakes présent dans des régions de contrôle. La région de signal pour le processus  $t\bar{t}H$  est typiquement une région avec au moins 5 jets et un  $b$ jets au moins (*cf* chapitre suivant). Nous conserverons toujours dans cette méthode la séparation entre les différents canaux  $e^\pm e^\pm$ ,  $e^\pm \mu^\pm$  et  $\mu^\pm \mu^\pm$ .

#### 4.3.1 Définition de la méthode

Pour cette estimation des événements Fakes nous allons diviser l'espace en 4 régions. Toutes les régions définies ont au moins un jet étiqueté  $b$ .

Ces quatre régions sont définies par :

Région de signal : au moins 5 jets. Nous la noterons RS par la suite.

Région de contrôle 1 : 2 ou 3 jets. Elle sera notée  $23TT$ .

Région de contrôle 2 : au moins 5 jets mais avec au moins un lepton qui sera défini comme **loose**. Elle sera notée  $5T\cancel{T}$ .

Région de contrôle 3 : 2 ou 3 jets avec au moins un lepton qui sera défini comme **loose**. Elle sera notée  $23T\cancel{T}$ .

Dans les régions  $23T\cancel{T}$  et  $5T\cancel{T}$ , le critère de mini-isolation (séparation minimum entre les traces correspondant au lepton et celles qui l'entourent) ne sera pas appliqué. Ces régions sont des régions enrichies en événements Fakes et sont considérées comme des régions de contrôle dans lesquelles il sera donc possible de regarder les données.

Nous considérons la grandeur  $\theta$  qui est défini comme étant le rapport entre le nombre d'événements dans la région 1 et ce nombre dans la région 3. Ce rapport étant indépendant du nombre de jets, il est possible d'extraire le nombre d'événements Fakes dans la région de signal à partir des trois autres régions (avec  $ll$  pour  $e^\pm e^\pm$ ,  $e^\pm \mu^\pm$  ou  $\mu^\pm \mu^\pm$ ) :

$$N_f^{RS})_l = \frac{N_f^{23TT} N_f^{5T\mathcal{T}'}}{N_f^{23T\mathcal{T}'}} = N_f^{5T\mathcal{T}'})_l \times \theta_l \quad (4.1)$$

Les nombres d'événements Fakes dans les deux régions à 2 et 3 jets et dans la région  $5T\mathcal{T}'$  seront mesurés grâce aux données après soustraction des événements de bruit de fond autres que les Fakes. Ceux-ci sont composés des bruits de fond irréductibles et des événements de Mis-Id. Ils sont obtenus par simulation Monte-Carlo, excepté le nombre d'événements Mis-Id dans la région  $23TT$  qui est estimé à partir des données (cf section 3.4.4).

$$N_f = N_{Données} - N_{MisId} - N_{Prompt} \quad (4.2)$$

### 4.3.2 Incertitude statistique

Dans cette méthode les erreurs statistiques ont pour origine le nombre d'événements de données dans les régions de contrôle 1,2 et 3. L'incertitude due à la taille finie des données est propagée pour obtenir une incertitude statistique sur le nombre d'événements Fakes dans la région de signal de l'ordre de 15% en  $e^\pm e^\pm$  et  $e^\pm \mu^\pm$  et 30% en  $\mu^\pm \mu^\pm$ .

### 4.3.3 Incertitudes systématiques

Les incertitudes systématiques sur le nombre d'événements fakes estimés auront deux origines principales :

- Incertitude systématique sur le nombre d'événements Mis-Id mesuré sur données pour les canaux  $e^\pm e^\pm$  et  $e^\pm \mu^\pm$  de l'ordre de 28% en  $e^\pm e^\pm$  et 14% en  $e^\pm \mu^\pm$ .
- Test de "fermeture" qui consiste à quantifier la validité de l'hypothèse de stabilité de  $\theta$  par rapport au nombre de jets. Ce test donne une incertitude de l'ordre de 33% en  $e^\pm e^\pm$ , 23% en  $e^\pm \mu^\pm$  et 52% en  $\mu^\pm \mu^\pm$ .

D'autres incertitudes peuvent être prises en compte telle que l'incertitude due aux nombres d'événements de bruits de fonds irréductibles dans la simulation Monte-Carlo. Mais celles-ci représentent une erreur sur le nombre d'événements Fakes estimés dans les régions de signal inférieure à 5% et est donc négligeable par rapport aux autres incertitudes. Nous n'en tiendrons donc pas compte.

### 4.3.4 Nombres d'événements *Fakes* obtenus

Grâce à cette méthode nous avons pu mesurer le nombre d'événements Fakes attendus dans la région de signal pour les trois canaux de notre étude. Ces valeurs sont exprimées dans le Tableau 4.5.

TAB. 4.5: Nombre d'événements <i>Fakes</i> dans la région de signal	
	$N_f$
$e^\pm e^\pm$	$7.66 \pm 1.14(Stat) \pm 2.54(fermeture) \pm 2.13(Mis - Id)$
$e^\pm \mu^\pm$	$12.96 \pm 1.97(stat) \pm 2.95(fermeture) \pm 1.75(Mis - Id)$
$\mu^\pm \mu^\pm$	$3.93 \pm 1.14(stat) \pm 2.04(fermeture)$

Cette méthode n'est pas la seule qui puisse être utilisée pour obtenir le nombre d'événements Fakes à partir des données. Les valeurs obtenues à partir d'une autre méthode de calcul à partir des données, dite

TAB. 4.6: Nombre d'événements *Fakes* dans la région de signal obtenus par la méthode de la matrice (MM) ou à partir des simulations Monte-Carlo

	$N_f^{MM}$	$N_f^{MC}$
$e^\pm e^\pm$	$10.49 \pm 1.93 \pm 3.15$	$3.88 \pm 1.10 \pm 0.38$
$e^\pm \mu^\pm$	$17.34 \pm 2.43 \pm 5.20$	$9.04 \pm 2.07 \pm 0.79$
$\mu^\pm \mu^\pm$	$2.18 \pm 0.66 \pm 0.66$	$3.19 \pm 1.06 \pm 0.31$

Méthode de la Matrice qui est détaillée dans la référence [10], et à partir des simulations Monte-Carlo sont indiqués dans le tableau 4.6.

La simulation Monte-Carlo fournit des valeurs inférieures aux deux autres méthodes pour les canaux  $e^\pm e^\pm$  et  $e^\pm \mu^\pm$ . Elle est incompatible avec les valeurs obtenues par la Méthode de la Matrice dans ceux-ci. Les valeurs fournies par la méthode présentée dans ce chapitre et celles obtenues par la Méthode de la Matrice sont compatibles pour les trois canaux.

# 5 Résultat et calcul de limite

## 5.1 Définition de la région de signal

Une région de signal est définie à partir de sélection sur des variables discriminantes. Celles qui sont utilisées sont :

- le nombre de jets (noté  $N_j$ ).
- le nombre de jets étiquetés  $b$  (noté  $N_b$ ).
- la somme des impulsions transverses des leptons et des jets, notés  $H_T$ .
- les impulsions transverses des leptons et des jets.
- la pseudorapidité  $\eta$  des leptons.
- la masse invariante des deux leptons (noté  $m_{ll}$ ).

En effectuant des sélections selon ces observables le but est de maximiser la présence du signal par rapport au bruit de fond. Le choix de ces variables a été motivé par le fait que le bruit de fond et le signal ne sont pas dans les mêmes régions pour celles-ci (comme indiqué dans le Tableau 5.1).

TAB. 5.1: Région de présence majoritaire pour différents bruits de fond et le signal selon trois variables discriminantes.

	$t\bar{t}V$	$VV$	$t\bar{t}(\text{Fakes/Mis-Id})$	$Z + jets(\text{Fakes/Mis-Id})$	$t\bar{t}H$
$N_j$	$\simeq 4$	$\leq 4$	$\simeq 4$	$\leq 4$	$\geq 4$
$N_b$	$\geq 1$	$= 0$	$\geq 1$	$= 0$	$\geq 1$
$m_{ll}$	–	$\simeq 91 \text{ GeV}$	–	–	–

Les sélections que nous allons effectuer porteront plus particulièrement sur le nombre de jets et le nombre de jets étiquetés  $b$  (voir la Figure 5.1). La distribution en  $H_T$  est très corrélée à ces deux grandeurs et il n'est donc pas nécessaire d'effectuer de coupures dessus.

Pour la suite de l'étude nous devons définir une région de signal où la signification statistique est maximale. Pour faire cela nous balayons les sélections possibles en nombre de jets et nombre de bjets pour choisir celle qui implique une signification statistique maximale (comme indiqué sur la Figure 5.2). La signification statistique est calculée d'après la formule :

$$\sigma_s = \frac{s}{\sqrt{b}} \quad (5.1)$$

où  $s$  est le nombre d'événements de signal et  $b$  celui de bruit de fond.

Finalement la région de signal choisie est définie par une sélection avec au moins 5 jets et au moins 1 jet étiqueté  $b$ .

La région de signal est définie par au moins 5 jets et 1 jet étiqueté  $b$ . La composition de cette région est donnée par canal dans le Tableau 5.2.

Il est maintenant nécessaire de savoir si notre sélection est suffisante, c'est-à-dire si elle nous permettra de

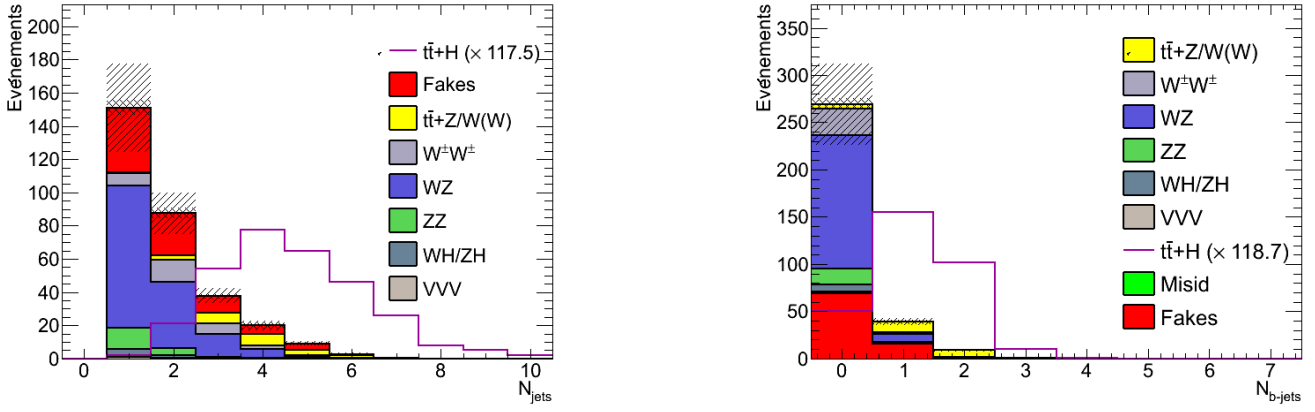


FIG. 5.1: Ces histogrammes présentent la distribution de deux variables discriminantes dans le canal  $\mu^+\mu^-$  pour des événements de signal et différents bruits de fond. Dans les deux cas la section efficace du signal est choisie de telle façon que l'intégrale du signal soit égale à l'intégrale de l'ensemble des bruits de fonds afin de le rendre visible. Le signal est représenté sur les graphes par une ligne continue.

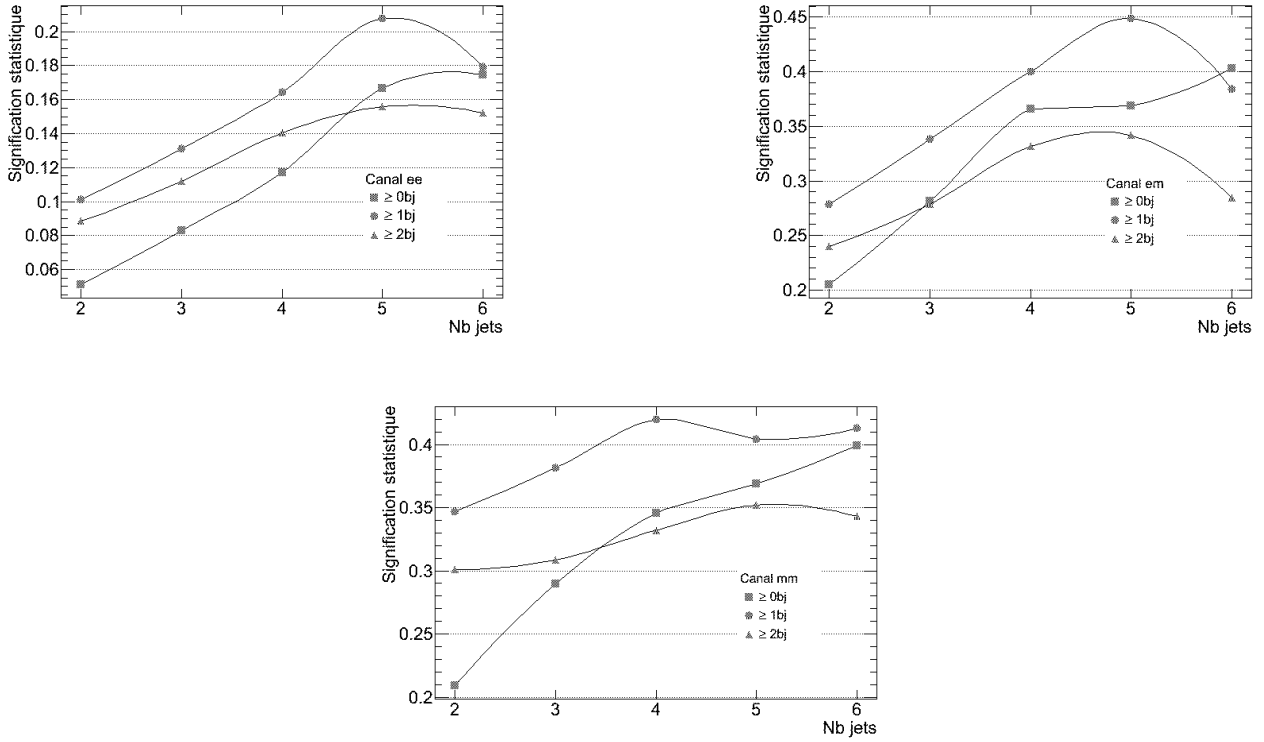


FIG. 5.2: Signification statistique en fonction du nombre de jets et du nombre de jets étiquetés dans les trois canaux.



discriminer notre bruit de fond de notre signal et si on pourra exclure notre hypothèse de signal dans le cas d'une observation des données correspondant à l'hypothèse du bruit de fond seul. Pour cela une limite est calculée en utilisant la méthode  $CL_s$ [11].

TAB. 5.2: Nombre d'événements de bruits de fond et de signal attendus dans la région de signal

	Bdf irréductible	Fakes	Misid	$t\bar{t}H$
$e^\pm e^\pm$	$2.65 \pm 0.14 \pm 0.34$	$7.66 \pm 1.14 \pm 3.32$	$3.59 \pm 0.15 \pm 1.08$	$0.68 \pm 0.06$
$e^\pm \mu^\pm$	$8.11 \pm 0.23 \pm 1.06$	$12.96 \pm 1.97 \pm 3.43$	$2.82 \pm 0.10 \pm 0.85$	$2.00 \pm 0.10$
$\mu^\pm \mu^\pm$	$5.30 \pm 0.18 \pm 0.78$	$3.93 \pm 1.14 \pm 2.04$	—	$1.18 \pm 0.08$

## 5.2 Définition de la méthode $CL_s$

### 5.2.1 Définition de la grandeur $CL_s$

Tout d'abord nous définissons deux hypothèses :

- une hypothèse de bruit de fond seul, notée "b".
- une hypothèse de bruit de fond plus signal, notée "s+b".

Pour ces deux hypothèses nous traçons leur fonction de probabilité en fonction du nombre d'événements. Dans les deux cas, les lois de probabilité suivies sont des lois de Poisson. Ensuite, pour un nombre d'événements observés donné, nous pouvons calculer pour chaque hypothèse une  $p$ -value<sup>1</sup>. Ainsi nous définissons  $CL_{s+b}$  comme étant la  $p$ -value dans le cas de l'hypothèse  $s+b$  et  $CL_b$  comme étant 1- $p$ -value pour l'hypothèse  $b$ . Ces différentes grandeurs sont illustrées sur la figure 5.3.

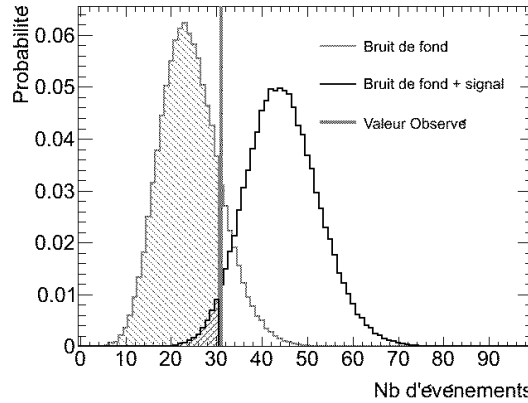


FIG. 5.3: Fonction de probabilité pour l'hypothèse  $b$  pour l'hypothèse  $s+b$ .  $CL_{s+b}$  et  $CL_b$  sont représentés par les surfaces respectivement hachurée vers la gauche et vers la droite.

Suite à ces définitions nous définissons aussi le rapport entre ces deux grandeurs que nous appelons  $CL_s$ .

$$CL_s = \frac{CL_{s+b}}{CL_b} \quad (5.2)$$

<sup>1</sup>Dans le cas d'un test statistique une  $p$ -value est la probabilité d'obtenir une valeur au moins aussi extrême que celle observée.

### 5.2.2 Limite observée

$CL_s$  quantifie la compatibilité entre une hypothèse s+b et les données. Ainsi on considère que l'on peut exclure un signal si lors d'une observation  $CL_s$  est égal (ou inférieure) à 5%. Dans ce cas notre signal est exclu avec un niveau de confiance de 95% (ou plus).

En observant les données on peut ainsi définir une limite, notée par la suite  $N_{up}$ , sur le nombre d'événements de signal présents. Pour cela on fait varier ce nombre jusqu'à obtenir  $CL_s = 0.05$  sous une hypothèse de nombre de bruit de fond fixe. On pourra alors exclure le signal que l'on cherche si le nombre d'événements attendu pour celui-ci est supérieur à cette limite observée.

### 5.2.3 Définition de la force du signal $\mu$

La force du signal peut être définie comme étant le rapport entre un nombre d'événements donné et le nombre d'événement nominal attendu pour le modèle étudié. Ainsi il est possible de définir une limite sur la force du signal à partir de la limite sur le nombre d'événement  $N_{up}$  :

$$\mu_{up} = \frac{N_{up}}{N_{nominal}} \quad (5.3)$$

Il est possible de calculer une limite directement sur la force du signal. La Figure 5.4 présente l'évolution de  $CL_s$  en fonction de la force du signal. La limite  $\mu_{up}$  correspond à la valeur de  $\mu$  telle que  $CL_s = 0.05$ .

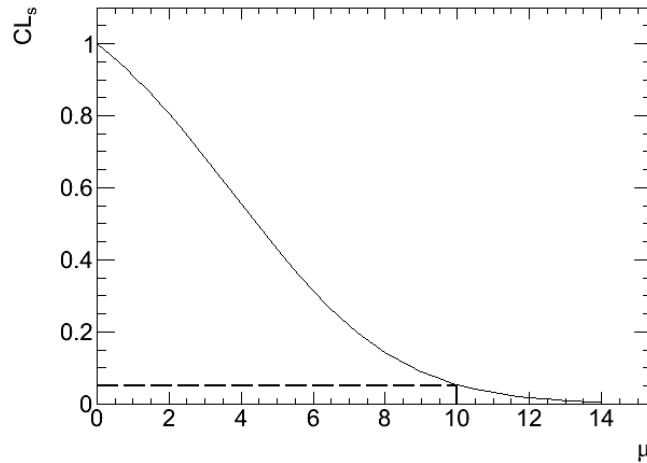


FIG. 5.4: Evolution de  $CL_s$  en fonction de force du signal et détermination de la valeur limite de force du signal

### 5.2.4 Limite attendue médiane

Par analogie avec une limite observée, il est possible de définir une limite attendue médiane à partir du nombre d'événements attendus dans l'hypothèse de bruit de fond. Pour cela on effectue la même opération que pour l'obtention de la limite observée en considérant cette fois-ci le nombre d'événements de bruit de fond médian comme étant la valeur observée dans les données. Il est aussi possible de définir des limites attendues à  $+1\sigma$ ,  $-1\sigma$ ,  $+2\sigma$ ,  $-2\sigma$ . Celles-ci sont obtenues en prenant le nombre d'événements médian plus (ou moins) le nombre de fois voulu de l'écart type comme valeur de nombre d'événements de données.

Le calcul de cette limite a deux buts :

- Juger de la sensibilité de notre étude. En effet, si la limite médiane attendue sur force du signal est très supérieur à un, alors notre analyse est peu sensible à notre signal. Le but d’une étude préliminaire à l’observation des données est de minimiser ce force du signal et si possible descendre aux alentours, voire en dessous de 1.
- Juger de la compatibilité entre les données et l’hypothèse de bruit de fond. Si les données sont disponibles, la comparaison des limites attendues (médiane, à  $\pm 1\sigma$  ou à  $\pm 2\sigma$ ) permettra de quantifier l’écart entre notre simulation des bruits de fond et les données. Par exemple si la limite observée est supérieure à la limite attendue à  $+3\sigma$  cela laisse penser qu’il y a du signal présent dans la région considérée.

### 5.3 Application à l’analyse du signal $t\bar{t}H$

Pour calculer la limite attendue médiane dans cette étude nous combinons les différents canaux. Les forces du signal limites que nous obtenons sont exprimées d’abord pour chaque canal séparément puis pour la combinaison des trois et les valeurs obtenues sont notées dans le Tableau 5.3.

TAB. 5.3: Limite attendue médiane sur  $\mu$  par canal puis pour la combinaison des trois canaux

<i>Canal <math>ee</math></i>	<i>Canal <math>\mu\mu</math></i>	<i>Canal <math>e\mu</math></i>	<i>Combinaison</i>
15.14	7.13	6.66	5.04

Finalement nous atteignons en combinant les trois canaux une force du signal de l’ordre de 5. Cela signifie que cette analyse est typiquement sensible au  $t\bar{t}H$  à condition que son amplitude soit cinq fois supérieure à celle qui est prédite par le Modèle Standard.

L’analyse peut donc être améliorée, par exemple, en modifiant la sélection des électrons (afin de réduire l’impact des bruits de fond instrumentaux) mais également combinée à d’autres signatures comme le canal à trois leptons.

# Conclusion

L'étude du couplage du boson de Higgs aux fermions est un terrain intéressant pour observer un indice de Nouvelle Physique si ce couplage n'est pas en adéquation avec celui prédit dans le Modèle Standard. Nous nous sommes ainsi intéressés au cours ce stage au couplage du boson de Higgs au quark top.

En premier lieu, le choix de la signature de notre signal, la présence de deux leptons de même signe dans l'état final, a été justifié. Suite à cela un choix de variables discriminantes a été fait pour maximiser la présence d'événements de signal par rapport au bruit de fond.

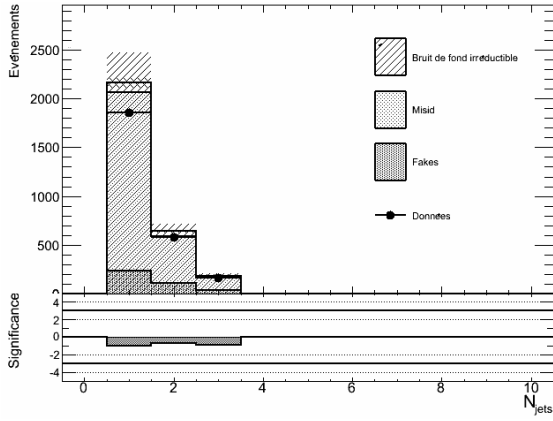
Dans un second temps, la validité de la simulation des bruits de fonds a été notre principale tâche. Dans ce cadre, une étude spécifique a été menée pour estimer à partir des données le nombre d'événements Fakes présents dans la région de signal.

Enfin nous avons regardé la sensibilité de notre analyse à notre signal par le calcul de la limite attendue médiane sur la force du signal pour la combinaison de nos trois canaux d'études. Nous pouvons en conclure que notre analyse n'est pas encore assez sensible à notre signal et qu'il faudrait faire des sélections supplémentaires pour augmenter la signification statistique.

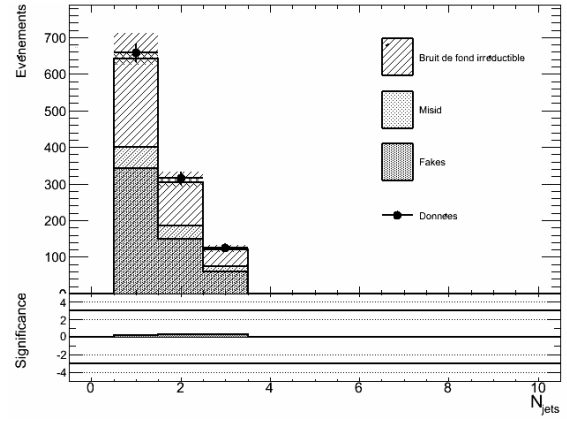
Une luminosité supplémentaire, comme celle attendue au RunII du LHC en 2015, permettra de gagner en précision. De plus, l'augmentation de l'énergie dans le référentiel du centre de masse de 8 TeV à 14 TeV entraînera une augmentation de la section efficace du signal d'un facteur 5 environ. Les bruits de fond ne devraient voir la leur n'augmenter que d'un facteur 2.

Le RunII du LHC offrira un contexte idéal pour la mesure de ce couplage du quark top au boson de Higgs, le plus fort du modèle Standard.

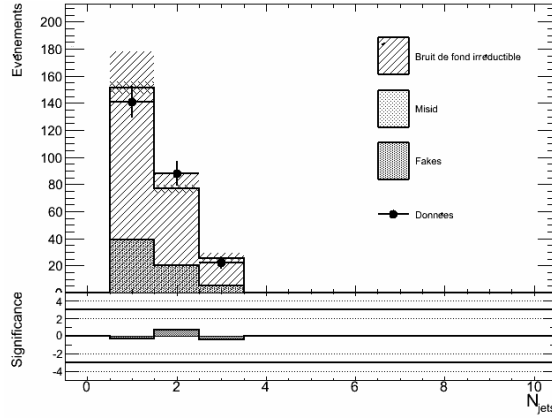
# A Distributions des bruits de fond et des données dans la régions de contrôle



(a) Canal  $e^+e^-$

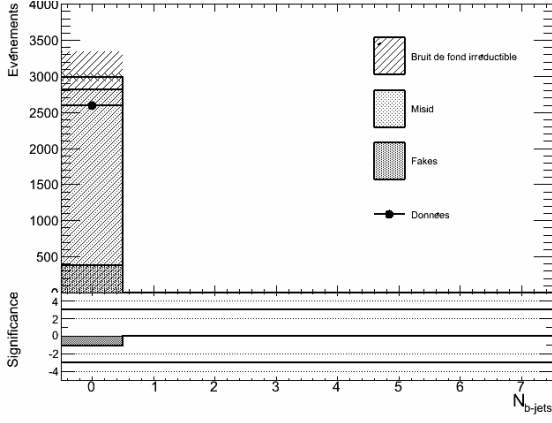


(b) Canal  $e^\pm\mu^\pm$

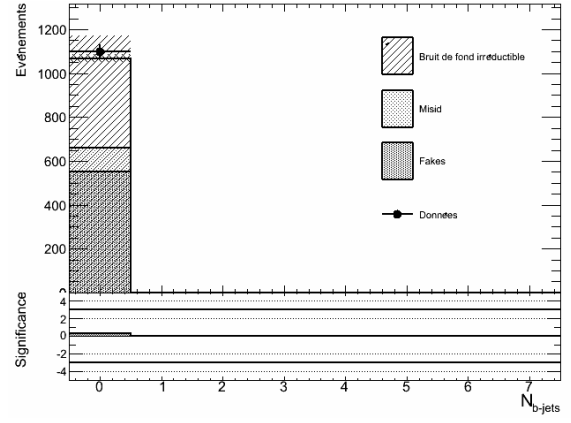


(c) Canal  $\mu^\pm\mu^\pm$

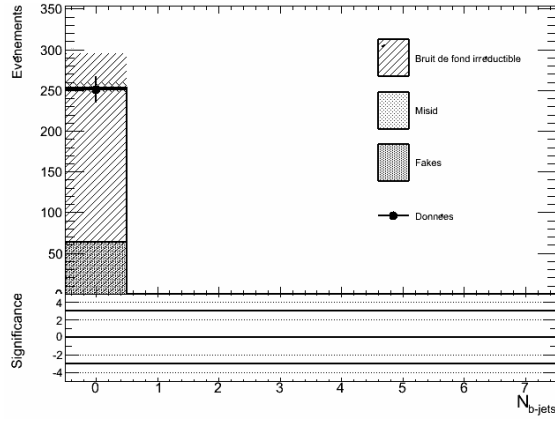
FIG. A.1: Distributions du nombre de jets dans les trois canaux pour le bruit de fond et les données dans la région de contrôle avec moins de 4 jets et aucun jets étiquetés  $b$ .



(a) Canal  $e^+e^-$

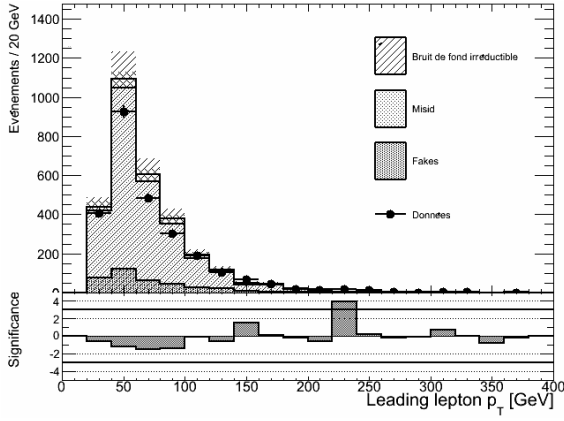


(b) Canal  $e^\pm\mu^\pm$

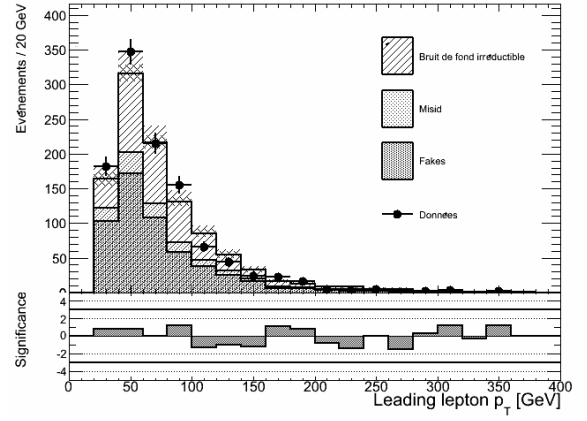


(c) Canal  $\mu^\pm\mu^\pm$

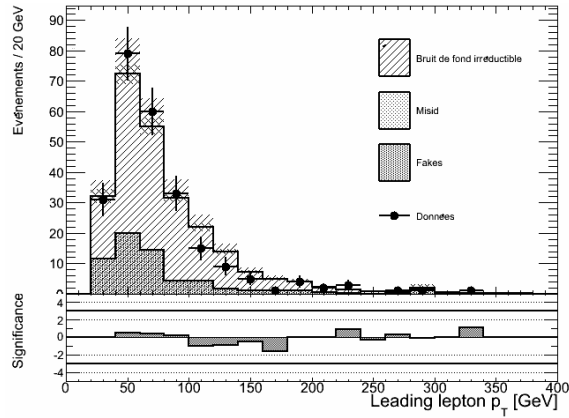
FIG. A.2: Distributions du nombre de  $b$ jets dans les trois canaux pour le bruit de fond et les données dans la région de contrôle avec moins de 4 jets et aucun jets étiquetés  $b$ .



(a) Canal  $e^{\pm}e^{\pm}$

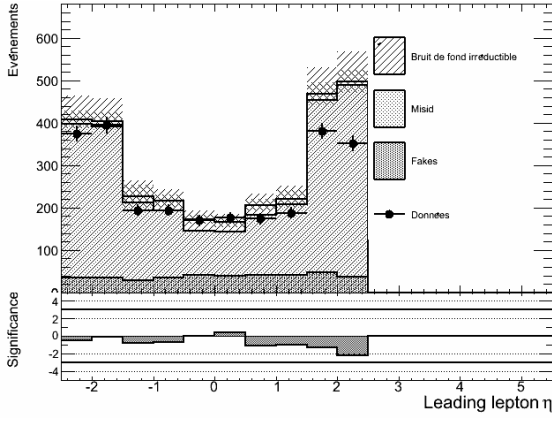


(b) Canal  $e^{\pm}\mu^{\pm}$

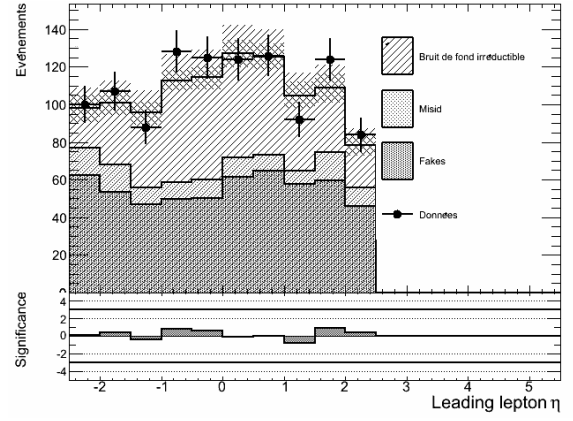


(c) Canal  $\mu^{\pm}\mu^{\pm}$

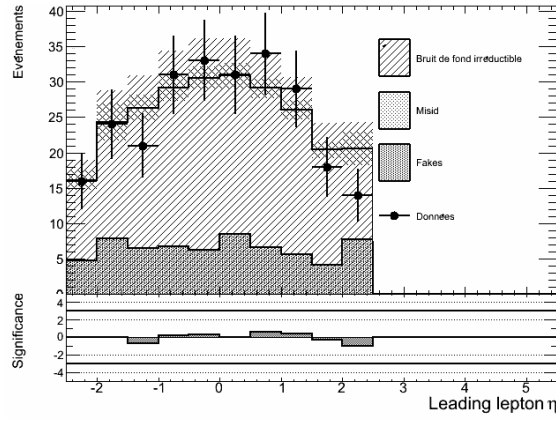
FIG. A.3: Distributions de l'impulsion transverse du lepton à plus haut  $p_T$  dans les trois canaux pour le bruit de fond et les données dans la région de contrôle avec moins de 4 jets et aucun jets étiquetés  $b$ .



(a) Canal  $e^{\pm}e^{\pm}$



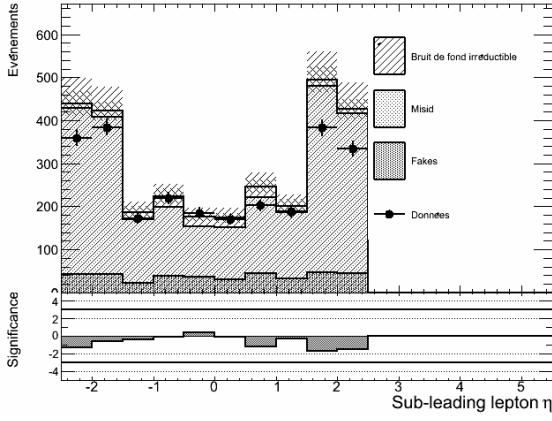
(b) Canal  $e^{\pm}\mu^{\pm}$



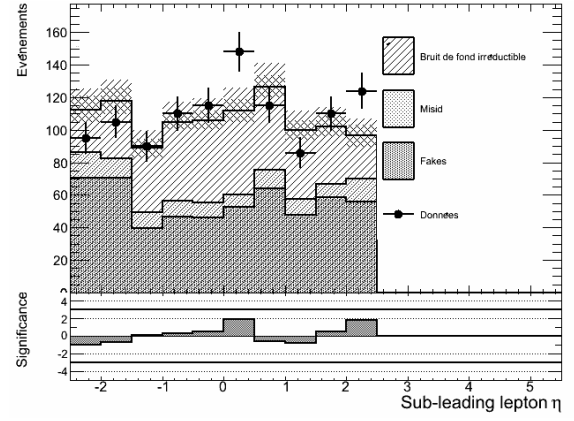
(c) Canal  $\mu^{\pm}\mu^{\pm}$

FIG. A.4: Distributions de la pseudorapidité du lepton à plus haut  $p_T$  dans les trois canaux pour le bruit de fond et les données dans la région de contrôle avec moins de 4 jets et aucun jets étiquetés  $b$ .

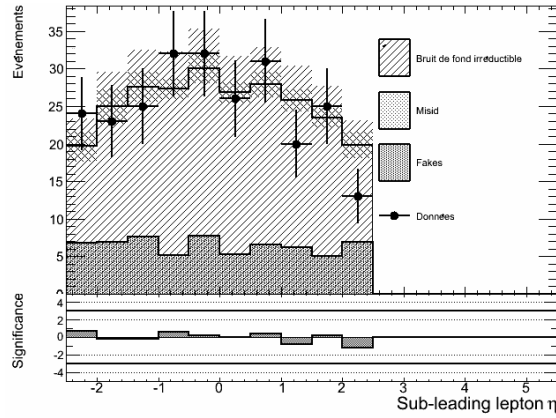




(a) Canal  $e^{\pm}e^{\pm}$



(b) Canal  $e^{\pm}\mu^{\pm}$



(c) Canal  $\mu^{\pm}\mu^{\pm}$

FIG. A.5: Distributions de la pseudorapidité du second lepton dans les trois canaux pour le bruit de fond et les données dans la région de contrôle avec moins de 4 jets et aucun jets étiquetés  $b$ .

# Bibliographie

- [1] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys.Lett.*, B716 :1–29, 2012.
- [2] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys.Lett.*, B716 :30–61, 2012.
- [3] J.Beringer et al. (particle data group). *Phys.Rev.*, D86 :010001, (2012) and 2013 partial update for the 2014 edition.
- [4] ATLAS Collaboration. Electron performance measurements with the atlas detector using the 2010 lh proton-proton collision data. *Eur. Phys. J. C72(2012)1909*, 2012.
- [5] <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections>.
- [6] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. MadGraph 5 : Going Beyond. *JHEP*, 1106 :128, 2011.
- [7] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 0605 :026, 2006.
- [8] Tanju Gleisberg, Frank Krauss, Andreas Schaliche, Steffen Schumann, and Jan-Christopher Winter. Studying  $W^+W^-$  production at the Fermilab Tevatron with SHERPA. *Phys.Rev.*, D72 :034028, 2005.
- [9] S. Agostinelli et al. GEANT4 : A Simulation toolkit. *Nucl.Instrum.Meth.*, A506 :250–303, 2003.
- [10] Search for anomalous production of events with same-sign dileptons and  $b$  jets in  $14.3 \text{ fb}^{-1}$  of  $pp$  collisions at  $\sqrt{s} = 8 \text{ TeV}$  with the ATLAS detector. Technical Report ATLAS-CONF-2013-051, CERN, Geneva, May 2013. Not published in the proceedings.
- [11] A L Read. Presentation of search results : the cl s technique. *Journal of Physics G : Nuclear and Particle Physics*, 28(10) :2693, 2002.

## Résumé

Le boson de Higgs a été découvert en Juillet 2012 au LHC (Large Hadron Collider), plus de 40 ans après que son existence ait été prédite. Cette découverte ouvre un nouveau champ de recherche pour l'étude des limites du Modèle Standard qui décrit notre monde à l'échelle élémentaire et la découverte de Nouvelle Physique grâce à l'étude des propriétés de ce boson.

Dans ce rapport, nous nous pencherons sur l'étude du couplage du boson de Higgs au fermion le plus massif, le quark top. Ce couplage est observable au LHC par l'étude de l'état final  $t\bar{t}H$ .

En premier lieu le Modèle Standard et la phénoménologie du boson de Higgs seront présentés. Le choix de deux leptons de même signe dans l'état final comme signature de notre signal, le processus  $t\bar{t}H$ , sera ensuite justifié.

Suite à ce choix de signature nous nous intéresserons à la simulation et à la validité de la modélisation des bruits de fond dans notre analyse. Une méthode d'estimation des événements Fakes dans la région de signal à partir des données sera introduite.

Enfin nous testerons la sensibilité de notre analyse à notre signal par un calcul de limite attendue sur force du signal par la méthode CLs.

## Abstract

The Higgs boson had been unobserved for more than 40 years after its existence was predicted. It was discovered in July 2012 at the LHC (Large Hadron Collider). This discovery offers many new way of testing the limits of the Standard Model which describes our world at the elementary scale. In fact the study of this boson propriety can lead to the observation of New Physic.

In this report we will deal with the study of the Higgs boson coupling to the top quark, the heaviest fermion, thanks to the  $t\bar{t}H$  final state.

Firstly we present the the Standard Model and the phenomenology of the Higgs boson. Thereafter we will justify the choice of looking at the  $t\bar{t}H$  with two same sign leptons signature

The next step will be to study the background processes in our study and to test their modelisation. A method to estimate Fakes events from the data in the signal region will be introduced.

Finally the sensitivity of our analysis to the signal will be quantified by the computation of an expected median limit on the signal strength with the CLs method.